

Mon Jan 13 10:35:08 2003

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 13, 2003, 09:53:15 ; Search time 31.9487 seconds
(without alignments)
58.391 Million cell updates/sec

Title: US-09-554-941-1
Perfect score: 70
Sequence: 1 VILGVPKIQFEG 14

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 908470 seqs, 133250620 residues 908470

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : 1: A.Geneseq.101002.*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
24: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the total score distribution.
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	70	100.0	14	20	AAV22918
2	48	68.6	79	22	AAAB80610
3	39	55.7	28	22	ABR31012
4	39	55.7	28	22	ABR36201
5	39	55.7	28	22	ABR21577
6	39	55.7	28	22	AAV56983
7	39	55.7	28	22	AAV69369
8	39	55.7	28	22	AAV17202
9	39	55.7	28	22	AAV29695
10	39	55.7	28	22	AAV04898

US-09-554-941-1.frag

11	39	55.7	28	23	ABG38985
12	39	55.7	285	21	AAV28710
13	39	55.7	508	21	AAV48146
14	39	55.7	508	23	ABR91938
15	39	55.7	510	23	ABR90889
16	39	55.7	533	21	AAV48145
17	39	55.7	564	21	AAV48144
18	39	55.7	627	22	ABP05180
19	39	54.3	70	23	ABP05116
20	38	54.3	100	22	AAV85222
21	38	54.3	616	22	AAV87613
22	38	54.3	616	22	AAV71827
23	38	54.3	616	22	AAV47509
24	37	52.9	321	23	AAE14458
25	37	52.9	347	22	ABG18656
26	37	52.9	385	21	AAV21893
27	37	52.9	392	20	AAV42969
28	37	52.9	406	21	AAV42968
29	37	52.9	415	22	AAV61116
30	37	52.9	471	21	AAV42967
31	37	52.9	478	21	AAV97293
32	37	52.9	564	21	AAV47029
33	37	52.9	566	19	AAV53900
34	37	52.9	566	21	AAV50004
35	37	52.9	965	22	AAV47030
36	37	52.9	1203	19	AAV47028
37	37	52.9	1211	19	AAV47028
38	37	52.9	1328	22	ABG23373
39	36	51.4	16	22	AAV97325
40	36	51.4	194	19	AAV30677
41	36	51.4	194	19	AAV60265
42	36	51.4	234	22	AAV65960
43	36	51.4	240	21	AAV07289
44	36	51.4	240	21	AAV07289
45	36	51.4	242	22	ABG05969

ALIGNMENTS

RESULT 1
ID AAV22918 standard: peptide: 14 AA.
XX AC AAV22918:
XX 19-AUG-1999 (first entry)
XX DE psmrta linker peptide used to make proteinase inhibitor fusions.
XX DE Peptide linker: fusion protein; pathogen resistance; pathogen tolerance;
XX KW plant; transgene; proteinase inhibitor; Arabidopsis thaliana;
XX KW parasitic nematode.
XX OS Synthetic.
XX PN WO928484-A1.
XX PD 10-JUN-1999.
XX PF 01-DEC-1998: 98WO-EP07792.
XX PR 03-DEC-1997: 97GB-0025556.
XX PA (NOVS) NOVARTIS AG.
XX PI Atkinson HJ, McPherson MJ, Urwin PE:
XX DR WPI: 1999-385387/32.
XX PT Proteinase inhibitor fusion proteins
XX PS Claim 10, page 23; 39pp; English.

Human peptide enco
Arabidopsis thaliana
Arabidopsis thaliana
Herbicidally activ
Herbicidally activ
Arabidopsis thaliana
Arabidopsis thaliana
Propionibacterium
Human OREX protein
Human Immune/Naema
ISAV structural pr
Infectious salmon
ISA antigen polype
Human protein diagno
Novel human diagno
Arabidopsis thaliana
Arabidopsis thaliana
Herbicidally activ
Herbicidally activ
Propionibacterium
Arabidopsis thaliana
Arabidopsis thaliana
Lipid associated p
Human N-proteinase
Amino acid sequenc
Bovine metallopro
Human N-proteinase
Novel human diagno
Novel human diagno
Le/SSB epitope pep
Glyceraldehyde pep
Klebsiella pneumon
Propionibacterium
Arabidopsis thaliana
Novel human diagno

xx The present sequence represents a peptide linker used in the fusion
 cc proteins of the invention. The specification describes a method for
 cc improving pathogen resistance or tolerance of a plant. The method
 cc comprises transformation of the plant with a transgene encoding a
 cc fusion protein which comprises two or more protein or domains that
 cc are capable of improving pathogen resistance or tolerance when
 cc expressed on their own. Specifically, two distinct proteinase
 cc inhibitors are co-delivered, as a fusion, to *Arabidopsis thaliana*. The
 cc used to improve pathogen resistance or tolerance of a plant and its
 cc descendants, especially against parasitic nematode attack.

Sequence 14 AA;
 Query Match
 Best Local Similarity 100.0%; Score 70; DB 20; Length 14;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 VILGVGPAKIOPEG 14
 Db 1 VILGVGPAKIOPEG 14

RESULT 2
 AAB80610
 ID AAB80610 standard; Protein; 79 AA.
 AC AAB80610;
 DT 02-MAY-2001 (first entry)
 DE Environmental stress tolerant protein SEQ ID 6.
 KW Environmental stress resistance; salt; heat; desert; transgenic plant.
 OS *Bruguiera seanguala*.
 PN WO200106006-A1.
 PD 25-JAN-2001.
 PF 19-JUL-2000; 2000WO-JP04862.
 PR 19-JUL-1999; 99JP-0235910.
 PR 24-MAR-2000; 2000JP-0085377.
 PA (NISC-) JAPAN SCI & TECHNOLOGY CORP.
 PI Yamada A, Ozeki Y, Saito T;
 DR WPI: 2001-147355/15.
 DR N-PSDB: AAF74189.
 XX Screening method to obtain DNA encoding environmental stress resistance
 PT factor, useful for producing transgenic plants resistant to
 PT environmental stress -
 PS Claim 19; Page 83-84; 167pp; Japanese.

Polynucleotide sequences AAF74187 - AAF74218 encode proteins
 cc AAB80608 - AAB80610, which impart environmental stress resistance
 cc Invention relates to a method for identifying DNA encoding proteins
 cc cDNA from a library originating in a salt-resistant organism into a host
 cc cell, culturing the transformants under conditions in which the
 cc untransformed host does not grow well, and selecting for viable clones.
 cc The method is useful for obtaining DNA encoding environmental stress
 cc resistance factors. The DNA encoding proteins conferring environmental
 cc stress resistance, can be used in the production of plants resistant to
 cc environmental stress, which can be cultivated in unfavourable
 cc environments such as deserts, salt damaged ground, cold regions and the
 cc oceans. They can be used for increasing the area of land covered by green

cc plants, and desert greening and afforestation, in order to counter the
 cc effects of the increase in atmospheric carbon dioxide concentration. PCR
 cc primers AAF74219 and AAF74220 are used in an example illustrating the
 cc method of the invention.

Sequence 79 AA;
 Query Match
 Best Local Similarity 68.6%; Score 48; DB 22; Length 79;
 Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
 QY 1 VILGVGPAKIOPEG 14
 Db 40 LVLGVPKRAHPEG 53

RESULT 3
 ABB31012
 ID ABB31012 standard; Peptide; 28 AA.
 AC ABB31012;
 DT 01-FEB-2002 (first entry)
 DE Peptide #3663 encoded by breast cell single exon nucleic acid probe.
 DE Human; microarray; single exon probe; gene expression; breast;
 KW disease; cancer.
 OS *Homo sapiens*.
 PN WO200157271-A2.
 PD 09-AUG-2001.
 PF 30-JAN-2001; 2001WO-US00662.
 PR 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 DR WPI: 2001-496933/54.

New spatially-addressable set of single exon nucleic acid probes,
 PT useful for measuring gene expression in sample derived from human
 PT breast, comprises number of single exon nucleic acid probes -
 PS Claim 27; SEQ ID NO 13980; 327pp + sequence listing; English.

The invention relates to a spatially-addressable set of single exon
 cc nucleic acid probes for measuring gene expression in a sample derived
 cc from human breast and BR 474 cells. The method involves contacting
 cc the probes with a collection of detectably labelled nucleic acids
 cc derived from mRNA of human breast, and then measuring the label
 cc bound to each probe of the microarray. The probes are useful for
 cc verifying the expression of regions of genomic DNA predicted to
 cc encode proteins. They are useful for gene discovery, and for
 cc determining predisposition and/or prognosis breast disease. Gene
 cc expression analysis is useful for assessing the toxicity of chemical
 cc agents on cells. The microarray of this invention presents a far greater
 cc diversity of probes for measuring gene expression, with far less bias
 cc than expressed sequence tag microarrays. The method is suitable for
 cc rapid production of functional information from genomic sequence. The
 cc present sequence is a peptide encoded by a single exon nucleic acid
 cc probe of the invention.

CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.

SO Sequence 28 AA;

Query Match 55.7%; Score 39; DB 22; Length 28;
Best Local Similarity 58.3%; Pred. No. 6.8;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 2 ILGVPKIOFE 13
: | | | | : | | |
Db 7 VLAVGPAQLQAE 18

RESULT 4
ABB36201
ID ABB36201 standard; Peptide; 28 AA.

AC ABB36201;

DE 04-FEB-2002 (first entry)

PE Peptide #3707 encoded by human foetal liver single exon probe.

Human; foetal liver; gene expression; single exon nucleic acid probe.

OS Homo sapiens.

PR WO200157277-A2.

PD 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US000669.

PR 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

PA (MOLE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;

PT WPI; 2001-483447/52.

PS Claim 27; SEQ ID NO 28836; 639pp + sequence listing; English.

XX The invention relates to a single exon nucleic acid probe for

XX measuring human gene expression in a sample derived from human foetal

XX liver. The single exon nucleic acid probes may be used for predicting,

XX measuring and displaying gene expression in samples derived from human

XX foetal liver. The present sequence is a peptide encoded by a single exon

XX nucleic acid probe of the invention.

CC Note: The sequence data for this patent did not form part of the

CC printed specification, but was obtained in electronic format directly

CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.

RESULT 5
ABB21577
ID ABB21577 standard; Protein; 28 AA.
AC ABB21577;
DE 23-JAN-2002 (first entry)
PE Protein #3576 encoded by probe for measuring heart cell gene expression.
Human; gene expression; heart; microarray; vascular system;
cardiovascular disease; hypertension; cardiac arrhythmia;
congenital heart disease.
OS Homo sapiens.
PR WO200157274-A2.
PD 09-AUG-2001.
PF 30-JAN-2001; 2001WO-US000666.
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
PA (MOLE-) MOLECULAR DYNAMICS INC.
PI Penn SG, Hanzel DK, Chen W, Rank DR;
PT WPI; 2001-488899/53.
PS Claim 15; SEQ ID NO 23347; 530pp; English.
XX Single exon nucleic acid probes for analyzing gene expression in human
XX hearts -
XX
XX The present invention relates to single exon nucleic acid probes for
XX measuring human gene expression in a sample derived from human heart (see
XX ABA21535-ABA41305). The present sequence is a protein encoded by one such
XX probe. The probes may be used for predicting, measuring and displaying
XX gene expression in samples derived from the human heart via microarrays.
XX By measuring gene expression, the probes are useful for predicting,
XX diagnosing, grading, staging, monitoring and prognosing diseases of the
XX human heart and vascular system e.g. cardiovascular disease,
XX hypertension, cardiac arrhythmias and congenital heart disease.
XX Note: The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.

SO Sequence 28 AA;

Query Match 55.7%; Score 39; DB 22; Length 28;
Best Local Similarity 58.3%; Pred. No. 6.8;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 2 ILGVPKIOFE 13
: | | | | : | | |
Db 7 VLAVGPAQLQAE 18

RESULT 6

AAM56983
ID AAM56983 standard; Protein; 28 AA.

AC AAM56983;

DT 05-NOV-2001 (first entry)

DE Human brain expressed single exon probe encoded protein SEQ ID NO: 29088.

KW Human: brain expressed exon; gene expression analysis; probe;
KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
KW epilepsy; cancer.

OS Homo sapiens.

PN W0200157275-A2.

PD 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US00667.

PR 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

PA (MOE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;

DR WPI: 2001-483446/52.

XX Single exon nucleic acid probes for analyzing gene expression in human

PT brains -

PS Example 4; SEQ ID NO: 29088; 650bp + Sequence Listing; English.

XX The present invention provides a number of single exon nucleic acid

CC probes which are derived from genomic sequences expressed in the human

CC brain. They can be used to measure gene expression in brain cell samples,

CC which may enable the diagnosis and improved treatment of nervous system

CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,

CC epilepsy and cancers. The present sequence is a protein encoded by one of

CC the probes of the invention.

SQ Sequence 28 AA:

Query Match

Best Local Similarity 55.7%; Score 39; DB 22; Length 28;

Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 ILGVGPAKIOFE 13

DB 7 VLAVGPAQLQAE 18

RESULT 7

AAM69369

ID AAM69369 standard; Protein; 28 AA.

AC AAM69369;

XX 06-NOV-2001 (first entry)

PF 30-JAN-2001; 2001WO-US00668.

PR 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

PA (MOE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;

DR WPI: 2001-488900/53.

XX Human genome-derived single exon nucleic acid probes useful for

PT analyzing gene expression in human bone marrow -

PS Example 4; SEQ ID NO: 29675; 658bp + Sequence Listing; English.

XX The present invention provides a number of single exon nucleic acid

CC probes which are derived from genomic sequences expressed in the human

CC bone marrow. They can be used to measure gene expression in bone marrow

CC samples, which may enable the improved diagnosis and treatment of cancers

CC such as lymphoma, leukemia and myeloma. The present sequence is a

CC protein encoded by one of the probes of the invention.

SQ Sequence 28 AA:

Query Match

Best Local Similarity 55.7%; Score 39; DB 22; Length 28;

Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 ILGVGPAKIOFE 13

DB 7 VLAVGPAQLQAE 18

RESULT 8

AAM17202

ID AAM17202 standard; Protein; 28 AA.

AC AAM17202;

XX 12-OCT-2001 (first entry)

DE Peptide #3636 encoded by probe for measuring cervical gene expression.

KW Probe; human; microarray; gene expression; cervical epithelial cell;

KW cervical cancer.

OS Homo sapiens.

PN W0200157278-A2.

PD 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US00670.

PR 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

PA (MOE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;

DR WPI: 2001-488901/53.

CC	from human placenta. The probes are useful for antenatal diagnosis of
CC	human genetic disorders.
XX	
SQ	Sequence 28 AA;
Query Match	55.7%; Score 39; DB 22; Length 28;
Best local Similarity	58.3%; Pred. No. 6.8;
Matches 7; Conservative	3; Mismatches 2; Indels 0; Gaps 0;
OY	2 ILGVPAKIOFE 13 : : Db 7 VLAAGPAAIQAE 18
RESULT 10	
AAM04898	
ID AAM04898	standard; Protein: 28 AA.
AC AAM04898;	
DT 09-OCT-2001	(first entry)
XX Peptide #3580 encoded by probe for measuring breast gene expression.	
DE Probe; human; breast disease; breast cancer; development disorder;	
XX inflammatory disease; proliferative breast disease; non-carcinoma tumour.	
KW Homo sapiens.	
XX MO2001S57270-AZ.	
PN 09-AUG-2001.	
PD 29-JAN-2001; 2001MO-US00661.	
PF	
XX 04-FEB-2000; 2000US-0180312.	
XX 26-MAY-2000; 2000US-0207456.	
PR 30-JUN-2000; 2000US-0608408.	
PR 03-AUG-2000; 2000US-0633666.	
PR 21-SEP-2000; 2000US-0234687.	
PR 27-SEP-2000; 2000US-0236359.	
PR 04-OCT-2000; 2000GB-0024263.	
XX PA (MOL-) MOLECULAR DYNAMICS INC.	
XX Penn SG, Hanzel DK, Chen W, Rank DR;	
PI WPI; 2001-476286/51.	
XX Novel single exon nucleic acid probe used to measuring gene expression	
PT in a human breast -	
PS Claim 27; SEQ ID NO 13638; 322pp; English.	
XX The present invention relates to novel single exon nucleic acid probes	
XX (see AAI00010-AAI10067). The present sequence is a peptide encoded by one	
CC such probe. The probes are useful for measuring human gene expression in	
CC a human breast sample, where the probe hybridises at high stringency to a	
CC nucleic acid expressed in the human breast. The probes are useful for	
CC predicting, diagnosing, grading, staging, monitoring and prognosing	
CC diseases of the human breast, particularly those diseases with polygenic	
CC aetiology. The diseases include: breast cancer, disorders of development,	
CC inflammatory diseases of the breast, fibrocystic changes, proliferative	
CC breast disease and non-carcinoma tumours.	
CC Note: The sequence data for this patent did not form part of the printed	
CC specification, but was obtained in electronic format directly from WIPO	
CC at ftp.wipo.int/pub/published_pct_sequences.	
XX SQ Sequence 28 AA;	
Query Match	55.7%; Score 39; DB 22; Length 28;
Best local Similarity	58.3%; Pred. No. 6.8;
Matches 7; Conservative	3; Mismatches 2; Indels 0; Gaps 0;

QY 2 ILGVGPAKIOFE 13
:1||||:1|
Db 7 VLAVGPAQLQAE 18

RESULT 11

ABG38985
ID ABG38985 standard; Peptide: 28 AA.

ABG38985;
XX

19-AUG-2002 (first entry)
XX

Human peptide encoded by genome-derived single exon probe SEQ ID 28650.
XX

Human: single exon probe; asthma; lung cancer; COPD; ILD;
XX chronic obstructive pulmonary disease; interstitial lung disease;
XX familial idiopathic pulmonary fibrosis; neurofibromatosis;
XX tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
XX Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
XX pulmonary histiocytosis; lymphangioleiomyomatosis; Karagener syndrome;
XX pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
XX primary ciliary dyskinesia; pulmonary hypertension;
XX hyaline membrane disease.
XX

Homo sapiens.
XX

WO200186003-A2.
XX

15-NOV-2001.
XX

30-JAN-2001; 2001WO-US00665.
XX

04-FEB-2000; 2000US-180312P.
XX 26-MAY-2000; 2000US-207456P.
XX 30-JUN-2000; 2000US-0608408.
XX 03-AUG-2000; 2000US-0632366.
XX 21-SEP-2000; 2000US-234687P.
XX 27-SEP-2000; 2000US-236359P.
XX 04-OCT-2000; 2000GB-0024263.
XX

(MOLE-) MOLECULAR DYNAMICS INC.
XX

Penn SG, Hanzel DK, Chen W, Rank DR;
XX

WPI; 2002-114183/15.
XX

Spatially-addressable set of single exon nucleic acid probes, used to
XX measure gene expression in human lung samples -
XX

Claim 27; SEQ ID No 28650; 634pp; English.
XX

The invention relates to a spatially-addressable set of single exon
XX nucleic acid probes for measuring gene expression in a sample derived
XX from human lung comprising single exon nucleic acid probes having one of
XX 12614 nucleic acid sequences mentioned in the specification, or their
XX complements or the 12387 open reading frames derived from the 12614
XX probes. Also included are a microarray comprising the novel set of
XX probes; the novel set of probes which hybridize at high stringency to a
XX nucleic acid expressed in the human lung; measuring gene expression in a
XX sample derived from human lung, comprising (a) contacting the array with
XX a collection of detectably labeled nucleic acids derived from human lung
XX mRNA, and (b) measuring the label detectably bound to each probe of
XX the array; identifying exons in a eukaryotic genome, comprising
XX of the eukaryote; and (b) detecting specific hybridisation of detectably
XX labeled nucleic acids from eukaryotic lung mRNA, to a single exon probe,
XX having a fragment identical to the predicted exon, the probe is included
XX in the above mentioned microarray; assigning exons to a single gene,
XX comprising (a) identifying exons from genomic sequence by the method
XX above and (b) measuring the expression of each of the exons in several
XX tissues and/or cell types using hybridisation to a single exon

CC microarrays having a probe with the exon, where a common pattern of
CC expression of the exons in the tissues and/or cell types indicates that
CC the exons should be assigned to a single gene; a peptide comprising one
CC of 12011 sequences, mentioned in the specification, or encoded by the
CC probes/open reading frames (ORF). The probes are used for gene
CC expression analysis, and for identifying exons in a gene, particularly
CC such as asthma, lung cancer, chronic obstructive pulmonary disease
CC (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary
CC fibrosis, neurofibromatosis, tuberous sclerosis, Gaucher's disease,
CC Niemann-Pick disease, Hermansky-Pudlak syndrome, sarcoidosis, pulmonary
CC haemosiderosis, pulmonary histiocytosis, lymphangioleiomyomatosis,
CC pulmonary alveolar proteinosis, Karagener syndrome, fibrocystic
CC and hyaline membrane disease, primary ciliary dyskinesia, pulmonary
CC encoded by a single exon probe of the invention.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences.

SO Sequence 28 AA.

Query Match 55.7%; Score 39; DB 23; Length 28;
Best Local Similarity 58.3%; Pred. No. 6.8;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 ILGVGPAKIOFE 13
:1||||:1|
Db 7 VLAVGPAQLQAE 18

RESULT 12

AAG28710
ID AAG28710 standard; Protein: 285 AA.

AAG28710;
XX

17-OCT-2000 (first entry)
XX

Arabidopsis thaliana protein fragment SEQ ID NO: 34031.
XX

Protein identification: signal transduction pathway; metabolic pathway;
XX hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.

Arabidopsis thaliana.
XX

EP103405-A2.
XX

06-SEP-2000.
XX

25-FEB-2000; 2000EP-0301439.
XX

25-FEB-1999; 99US-0121825.
XX 05-MAR-1999; 99US-0123180.
XX 09-MAR-1999; 99US-0123548.
XX 23-MAR-1999; 99US-0125788.
XX 25-MAR-1999; 99US-0126264.
XX 29-MAR-1999; 99US-0126785.
XX 01-APR-1999; 99US-0127462.
XX 06-APR-1999; 99US-0128234.
XX 08-APR-1999; 99US-0128714.
XX 16-APR-1999; 99US-0129845.
XX 19-APR-1999; 99US-0130077.
XX 21-APR-1999; 99US-0130449.
XX 23-APR-1999; 99US-0130510.
XX 28-APR-1999; 99US-0130891.
XX 30-APR-1999; 99US-0131449.
XX 30-APR-1999; 99US-0132048.
XX 04-MAY-1999; 99US-0132407.
XX 05-MAY-1999; 99US-0132484.
XX 05-MAY-1999; 99US-0132485.

PR 21-OCT-1999; 99US-0160814.
 PR 21-OCT-1999; 99US-0160815.
 PR 22-OCT-1999; 99US-0160980.
 PR 22-OCT-1999; 99US-0160981.
 PR 22-OCT-1999; 99US-0160981.
 PR 25-OCT-1999; 99US-0161404.
 PR 25-OCT-1999; 99US-0161405.
 PR 25-OCT-1999; 99US-0161406.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161921.
 PR 28-OCT-1999; 99US-0161922.
 PR 29-OCT-1999; 99US-0162142.

Query Match 55.7% Score 39; DB 21; Length 285;
 Best Local Similarity 42.9%; Pred. No. 81;
 Matches 6; Conservative 4; Mismatches 4; Indels 0; Gaps 0;
 OY 1 VIIGVPAKIOFEG 14
 Db 20 VVVGGPTEVFEG 33

RESULT 13
 AAG48146
 ID AAG48146 standard; Protein: 508 AA.
 AC AAG48146;
 XX
 DT 18-OCT-2000 (first entry)
 XX
 DE Arabidopsis thaliana protein fragment SEQ ID NO: 60770.

KW Protein identification: signal transduction pathway; metabolic pathway;
 KW hybridisation assay; genetic mapping; gene expression control; promoter;
 KW termination sequence.
 OS Arabidopsis thaliana.
 XX
 XX EPI033405-A2.
 PD 06-SEP-2000.
 XX
 XX 25-FEB-2000; 2000EP-0301439.

PR 25-FEB-1999; 99US-0121825.
 PR 05-MAR-1999; 99US-0123180.
 PR 09-MAR-1999; 99US-0123548.
 PR 23-MAR-1999; 99US-0125788.
 PR 25-MAR-1999; 99US-0126264.
 PR 29-MAR-1999; 99US-0126785.
 PR 01-APR-1999; 99US-0127462.
 PR 06-APR-1999; 99US-0128234.
 PR 08-APR-1999; 99US-0128714.
 PR 16-APR-1999; 99US-0129845.
 PR 19-APR-1999; 99US-0130077.
 PR 21-APR-1999; 99US-0130449.
 PR 23-APR-1999; 99US-0130510.
 PR 28-APR-1999; 99US-0130891.
 PR 30-APR-1999; 99US-0131449.
 PR 30-APR-1999; 99US-0132048.
 PR 04-MAY-1999; 99US-0132484.
 PR 05-MAY-1999; 99US-0132485.
 PR 06-MAY-1999; 99US-0132486.
 PR 07-MAY-1999; 99US-0132487.
 PR 11-MAY-1999; 99US-0132863.
 PR 14-MAY-1999; 99US-0134256.
 PR 14-MAY-1999; 99US-0134218.
 PR 14-MAY-1999; 99US-0134221.

PR 14-MAY-1999; 99US-0134370.
 PR 18-MAY-1999; 99US-0134768.
 PR 19-MAY-1999; 99US-0134941.
 PR 20-MAY-1999; 99US-0135124.
 PR 21-MAY-1999; 99US-0135333.
 PR 24-MAY-1999; 99US-0135629.
 PR 25-MAY-1999; 99US-0136021.
 PR 27-MAY-1999; 99US-0136392.
 PR 28-MAY-1999; 99US-0136782.
 PR 01-JUN-1999; 99US-0137222.
 PR 03-JUN-1999; 99US-0137528.
 PR 04-JUN-1999; 99US-0137528.
 PR 07-JUN-1999; 99US-0137528.
 PR 08-JUN-1999; 99US-0137724.
 PR 10-JUN-1999; 99US-0138094.
 PR 10-JUN-1999; 99US-0138540.
 PR 14-JUN-1999; 99US-0138847.
 PR 16-JUN-1999; 99US-0139119.
 PR 16-JUN-1999; 99US-0139452.
 PR 17-JUN-1999; 99US-0139453.
 PR 18-JUN-1999; 99US-0139452.
 PR 18-JUN-1999; 99US-0139454.
 PR 18-JUN-1999; 99US-0139455.
 PR 18-JUN-1999; 99US-0139456.
 PR 18-JUN-1999; 99US-0139457.
 PR 18-JUN-1999; 99US-0139458.
 PR 18-JUN-1999; 99US-0139459.
 PR 18-JUN-1999; 99US-0139460.
 PR 18-JUN-1999; 99US-0139461.
 PR 18-JUN-1999; 99US-0139462.
 PR 18-JUN-1999; 99US-0139463.
 PR 18-JUN-1999; 99US-0139750.
 PR 21-JUN-1999; 99US-0139763.
 PR 22-JUN-1999; 99US-0139817.
 PR 23-JUN-1999; 99US-0139859.
 PR 23-JUN-1999; 99US-0140353.
 PR 24-JUN-1999; 99US-0140354.
 PR 28-JUN-1999; 99US-0140695.
 PR 29-JUN-1999; 99US-0140823.
 PR 30-JUN-1999; 99US-0140991.
 PR 01-JUL-1999; 99US-0141287.
 PR 01-JUL-1999; 99US-0141842.
 PR 02-JUL-1999; 99US-0142154.
 PR 06-JUL-1999; 99US-0142055.
 PR 08-JUL-1999; 99US-0142390.
 PR 09-JUL-1999; 99US-0142803.
 PR 12-JUL-1999; 99US-0142920.
 PR 13-JUL-1999; 99US-0142977.
 PR 14-JUL-1999; 99US-0143624.
 PR 15-JUL-1999; 99US-0144005.
 PR 16-JUL-1999; 99US-0144085.
 PR 19-JUL-1999; 99US-0144086.
 PR 19-JUL-1999; 99US-0144325.
 PR 19-JUL-1999; 99US-0144331.
 PR 19-JUL-1999; 99US-0144332.
 PR 19-JUL-1999; 99US-0144333.
 PR 19-JUL-1999; 99US-0144334.
 PR 20-JUL-1999; 99US-0144335.
 PR 20-JUL-1999; 99US-0144352.
 PR 20-JUL-1999; 99US-0144632.
 PR 21-JUL-1999; 99US-0144684.
 PR 21-JUL-1999; 99US-0144814.
 PR 21-JUL-1999; 99US-0145086.
 PR 22-JUL-1999; 99US-0145088.
 PR 22-JUL-1999; 99US-0145089.
 PR 22-JUL-1999; 99US-0145089.
 PR 22-JUL-1999; 99US-0145192.
 PR 23-JUL-1999; 99US-0145145.
 PR 23-JUL-1999; 99US-0145218.
 PR 23-JUL-1999; 99US-0145224.
 PR 26-JUL-1999; 99US-0145276.
 PR 27-JUL-1999; 99US-0145913.

PR 27-JUL-1999; 99US-0145918.
 PR 27-JUL-1999; 99US-0145919.
 PR 28-JUL-1999; 99US-0145951.
 PR 02-AUG-1999; 99US-0146386.
 PR 02-AUG-1999; 99US-0146388.
 PR 02-AUG-1999; 99US-0146389.
 PR 03-AUG-1999; 99US-0147038.
 PR 04-AUG-1999; 99US-0147204.
 PR 04-AUG-1999; 99US-0147302.
 PR 05-AUG-1999; 99US-0147192.
 PR 06-AUG-1999; 99US-0147260.
 PR 06-AUG-1999; 99US-0147303.
 PR 06-AUG-1999; 99US-0147416.
 PR 09-AUG-1999; 99US-0147493.
 PR 09-AUG-1999; 99US-0147935.
 PR 10-AUG-1999; 99US-0148171.
 PR 11-AUG-1999; 99US-0148319.
 PR 12-AUG-1999; 99US-0148341.
 PR 13-AUG-1999; 99US-0148565.
 PR 13-AUG-1999; 99US-0148684.
 PR 16-AUG-1999; 99US-0149368.
 PR 17-AUG-1999; 99US-0149426.
 PR 18-AUG-1999; 99US-0149426.
 PR 20-AUG-1999; 99US-0149722.
 PR 20-AUG-1999; 99US-0149723.
 PR 20-AUG-1999; 99US-0149929.
 PR 23-AUG-1999; 99US-0149902.
 PR 23-AUG-1999; 99US-0149930.
 PR 25-AUG-1999; 99US-0150566.
 PR 26-AUG-1999; 99US-0150884.
 PR 27-AUG-1999; 99US-0151065.
 PR 27-AUG-1999; 99US-0151066.
 PR 27-AUG-1999; 99US-0151080.
 PR 30-AUG-1999; 99US-0151303.
 PR 31-AUG-1999; 99US-0151348.
 PR 01-SEP-1999; 99US-0151330.
 PR 07-SEP-1999; 99US-0152363.
 PR 10-SEP-1999; 99US-0153070.
 PR 13-SEP-1999; 99US-0153758.
 PR 15-SEP-1999; 99US-0154018.
 PR 16-SEP-1999; 99US-0154039.
 PR 20-SEP-1999; 99US-0154779.
 PR 22-SEP-1999; 99US-0155139.
 PR 23-SEP-1999; 99US-0155486.
 PR 24-SEP-1999; 99US-0155659.
 PR 28-SEP-1999; 99US-0156458.
 PR 29-SEP-1999; 99US-0156596.
 PR 04-OCT-1999; 99US-0157117.
 PR 05-OCT-1999; 99US-0157753.
 PR 06-OCT-1999; 99US-0157865.
 PR 07-OCT-1999; 99US-0158029.
 PR 08-OCT-1999; 99US-0158232.
 PR 12-OCT-1999; 99US-0158369.
 PR 13-OCT-1999; 99US-0159293.
 PR 13-OCT-1999; 99US-0159294.
 PR 13-OCT-1999; 99US-0159329.
 PR 14-OCT-1999; 99US-0159325.
 PR 14-OCT-1999; 99US-0159330.
 PR 14-OCT-1999; 99US-0159331.
 PR 14-OCT-1999; 99US-0159637.
 PR 14-OCT-1999; 99US-0159637.
 PR 14-OCT-1999; 99US-0159638.
 PR 18-OCT-1999; 99US-0159584.
 PR 21-OCT-1999; 99US-0160741.
 PR 21-OCT-1999; 99US-0160767.
 PR 21-OCT-1999; 99US-0160768.
 PR 21-OCT-1999; 99US-0160770.
 PR 21-OCT-1999; 99US-0160814.
 PR 21-OCT-1999; 99US-0160815.
 PR 22-OCT-1999; 99US-0160980.
 PR 22-OCT-1999; 99US-0160981.
 PR 22-OCT-1999; 99US-0160989.
 PR 23-OCT-1999; 99US-0161404.
 PR 25-OCT-1999; 99US-0161405.

PR 25-OCT-1999; 99US-0161406.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161992.
 PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.

Query Match 55.7%; Score 39; DB 21; Length 508;
 Best Local Similarity 42.9%; Pred. No. 1.5e+02;
 Matches 6; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Oy 1 VILVGPAKIOPEG 14
 Db 243 VVVGSGPTGVFESG 256

RESULT 14

ABB91938 standard; Protein: 508 AA.

AC ABB91938;

DT 31-MAY-2002 (first entry)

DE Herbicidally active polypeptide SEQ ID NO 1149.

Herbicidal; plant; agriculture; herbicide.

OS Arabidopsis thaliana.

PN WO200210210-A2.

PD 07-FEB-2002.

PF 28-AUG-2001; 2001WO-EP09892.

PR 28-AUG-2001; 2001WO-EP09892.

PA (FARB) BAYER AG.

PI Tietjen K, Weidner M;

DR WPI, 2002-269010/31.

XX Identifying plant target proteins for herbicidally active compounds,
 XX comprising aligning and comparing nucleic acid or amino acid sequences
 XX from plant with nucleic acid or amino acid sequences from non-plant
 XX organisms -

PS Claim 5; SEQ ID NO 1149; 261pp + Sequence listing; English.

XX The invention relates to identifying target proteins
 CC (ABB90790-ABB94016) for herbicidally active compounds, comprising
 CC aligning and comparing nucleic acid or amino acid sequences from plant
 CC with nucleic acid or amino acid sequences from non-plant organisms using
 CC suitable search parameters, where plant sequences having an E-value
 CC greater by a factor of 3 than the E-value of most similar non-plant
 CC sequences are selected. The polypeptides or nucleic acids encoding them
 CC are useful for identifying modulators. The identified modulators are
 CC useful as herbicides.

CC Sequence 508 AA;

Query Match 55.7%; Score 39; DB 23; Length 508;

Best Local Similarity 42.9%; Pred. No. 1.5e+02;

Matches 6; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Oy 1 VILVGPAKIOPEG 14
 Db 243 VVVGSGPTGVFESG 256

```

RESULT 15
ABB90889
ID ABB90889 standard; Protein: 510 AA.
XX
AC ABB90889;
XX
DT 31-MAY-2002 (first entry)
XX
DE Herbicidally active polypeptide SEQ ID NO 100.
XX
KW Herbicidal; plant; agriculture; herbicide.
XX
OS Arabidopsis thaliana.
XX
PN MO200210210-A2.
XX
PD 07-FEB-2002.
XX
PF 28-AUG-2001; 2001MO-EP09892.
XX
PR 28-AUG-2001; 2001MO-EP09892.
XX
PA (FARB ) BAYER AG.
XX
PI Tietjen K, Weidler M;
XX
DR WPI; 2002-269010/31.
XX
PT Identifying plant target proteins for herbicidally active compounds,
PT comprising aligning and comparing nucleic acid or amino acid sequences
PT from plant with nucleic acid or amino acid sequences from non-plant
PT organisms -
XX
PS Claim 5; SEQ ID NO 100; 261pp + Sequence Listing; English.
XX
CC The invention relates to identifying target proteins
CC (ABB90790-ABB94016) for herbicidally active compounds, comprising
CC aligning and comparing nucleic acid or amino acid sequences from plant
CC with nucleic acid or amino acid sequences from non-plant organisms using
CC suitable search parameters, where plant sequences having an E-value
CC greater by a factor of 3 than the E-value of most similar non-plant
CC sequences are selected. The polypeptides or nucleic acids encoding them
CC are useful for identifying modulators. The identified modulators are
CC useful as herbicides.
XX
SQ Sequence 510 AA;

```

Query Match 55.78; Score 39; DB 23; Length 510;
 Best Local Similarity 42.9%; Pred. No. 1.5e+02;
 Matches 6; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

```

OY 1 VIIVGPGAKIOPEG 14
   |::|||::||
DB 245 VVVGPGTGVFESG 258

```

Search completed: January 13, 2003, 09:55:40
 Job time : 33.9487 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 13, 2003, 09:53:20 ; Search time 12.9231 Seconds
(without alignments)
104.146 Million cell updates/sec

Title: US-09-554-941-1
Perfect score: 70
Sequence: 1 VILGVGPAKIOFEG 14

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 segs, 96134422 residues
Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	70	100.0	75	2	T12188
2	70	100.0	75	2	S09098
3	64	91.4	75	2	S37240
4	56	80.0	77	2	T12187
5	53	75.7	74	1	SMBH1
6	53	75.7	77	2	S37239
7	49	70.0	79	2	J02128
8	43	61.4	80	2	T10087
9	42	60.0	78	2	S48038
10	40	57.1	84	2	P70569
11	40	57.1	82	2	T07076
12	40	57.1	82	2	T07104
13	40	57.1	212	2	P91002
14	40	57.1	215	2	E85847
15	40	57.1	456	2	AE0164
16	39	55.7	220	1	B5AG58
17	39	55.7	220	1	AI3248
18	39	55.7	360	2	C84413
19	39	55.7	508	2	T02486
20	39	55.7	512	2	H86206
21	39	55.7	813	2	H86206
22	38	54.3	77	2	S52636
23	38	54.3	82	2	T03727
24	38	54.3	237	2	D71199
25	38	54.3	249	2	A48325
26	38	54.3	334	2	G69303
27	38	54.3	390	2	S75876
28	38	54.3	430	2	C97160
29	38	54.3	915	2	S44797

30	38	54.3	2632	2	T18718	dynein heavy chain
31	37	52.9	248	2	A70466	transcription anti
32	37	52.9	304	2	S13533	mRNA splice defec
33	37	52.9	314	2	S55179	mRNA splice defec
34	37	52.9	334	2	A36707	protein-glutamate
35	37	52.9	350	2	AD3560	methylated-DNA-lpr
36	37	52.9	371	2	T49100	hypothetical prote
37	37	52.9	375	1	S66272	alcohol dehydrogen
38	37	52.9	392	2	T51772	acetyl-CoA C-acety
39	37	52.9	406	2	T05865	hypothetical prote
40	37	52.9	421	2	A12824	NADH dehydrogenase
41	37	52.9	438	2	H97602	probable NADH dehy
42	37	52.9	447	2	T09414	pectinesterase hom
43	37	52.9	458	1	C70058	conserved hypothet
44	37	52.9	508	2	H90242	glycine dehydrogen
45	37	52.9	554	2	B85072	hypothetical prote

ALIGNMENTS

RESULT 1

T12188
metallothionein - fava bean
C:Species: Vicia faba (fava bean)
C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 21-Jul-2000
C:Accession: T12188
R:Foley, R.C.; Liang, Z.M.; Singh, K.B.
Plant Mol. Biol. 33, 583-591, 1997
A>Title: Analysis of type 1 metallothionein cDNAs in Vicia faba.
A:Reference number: 217444; MUID:97238465; PMID:9132050
A:Accession: T12188
A:status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-75 <POL>
A:Cross-references: EMBL:X91078; NID:91150656; PIDN:CAA62552.1; PID:91150657
A:Experimental source: leave
A:Genetics: mlb
C:Superfamily: metallothionein
C:Keywords: metal binding

Query Match 100.0%; Score 70; DB 2; Length 75;
Best Local Similarity 100.0%; Pred. No. 1.5e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VILGVGPAKIOFEG 14
DB 36 VILGVGPAKIOFEG 49

RESULT 2

S09098
metallothionein - garden pea
C:Species: Pisum sativum (garden pea)
C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 20-Aug-1999
C:Accession: S09098; S20347; S60719
R:Evans, I.M.; Gatehouse, L.N.; Gatehouse, J.A.; Robinson, N.J.; Croy, R.R.D.
FEBS Lett. 262, 29-32, 1990
A>Title: A gene from pea (Pisum sativum L.) with homology to metallothionein genes.
A:Reference number: S09098; MUID:90201367; PMID:2318309
A:Accession: S09098
A:Molecule type: DNA
A:Residues: 1-75 <EVA>
A:Cross-references: EMBL:Z23097; NID:9312504; PIDN:CAA80645.1; PID:9312505
R:Kille, P.; Winge, D.R.; Hatwood, J.L.; Kay, J.
FEBS Lett. 295, 171-175, 1991
A>Title: A plant metallothionein produced in E. coli.
A:Reference number: S20347; MUID:92111733; PMID:1765150
A:Accession: S20347
A:Molecule type: mRNA
A:Residues: 1-75 <KIT>
A:Cross-references: GB:Z23097; NID:9312504; PIDN:CAA80645.1; PID:9312505

A:Experimental source: cv. Feltham First
 A:Accession: S60719
 A:Molecule type: protein
 A:Residues: 2-4;26-28;39-42;44-45;47-48;54-55 <K1W>
 A:Experimental source: cv. Feltham First
 C:Genetics: 17/2
 A:Introns: 17/2
 C:Superfamily: metallochionein
 C:Keywords: metal binding

Query Match
 Best Local Similarity 100.0%; Score 70; DB 2; Length 75;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VILGVPKIOPEG 14
 |||||
 Db 36 VILGVPKIOPEG 49

RESULT 3

S37240
 metallochionein-like protein - white clover
 C:Species: Trifolium repens (white clover)
 C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 20-Aug-1999
 C:Accession: S37240
 R:Ellison, N.W.
 submitted to the EMBL Data Library, September 1993
 A:Description: Sequence analysis of two cDNA clones for metallochionein-like proteins fr
 A:Reference number: S37239
 A:Accession: S37240
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-75 <ELL>
 A:Cross-references: EMBL:226493; NID:9403328; PIDN:CAA81265.1; PID:9403329
 C:Superfamily: metallochionein

Query Match
 Best Local Similarity 91.4%; Score 64; DB 2; Length 75;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VILGVPKIOPEG 13
 |||||
 Db 36 VILGVPKIOPEG 48

RESULT 4

T12187
 metallochionein, type 1 - fava bean
 C:Species: Vicia faba (fava bean)
 C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 21-Jul-2000
 C:Accession: T12187
 R:Pooley, R.C.; Liang, Z.M.; Singh, K.B.
 Plant Mol. Biol. 33, 583-591, 1997
 A:Title: Analysis of type 1 metallochionein cDNAs in Vicia faba.
 A:Reference number: Z17444; MUID:97238465; PMID:9132050
 A:Accession: T12187
 A:Status: preliminary; translated from GB/EMBL/DBD
 A:Molecule type: mRNA
 A:Residues: 1-77 <POL>
 A:Cross-references: EMBL:X91077; NID:91150654; PIDN:CAA62551.1; PID:91150655
 A:Experimental source: leaves
 C:Superfamily: metallochionein
 C:Keywords: metal binding

Query Match
 Best Local Similarity 80.0%; Score 56; DB 2; Length 77;
 Matches 10; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 VILGVPKIOPEG 14
 |||||
 Db 39 VILGVPKIOPEG 52

RESULT 5

SMBH1
 metallochionein - barley
 C:Species: Hordeum vulgare (barley)
 C:Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 22-Jun-1999
 C:Accession: S17299; S15558
 R:Okumura, N.; Nishizawa, N.K.; Uehara, Y.; Mori, S.
 Plant Mol. Biol. 17, 531-533, 1991
 A:Title: An iron deficiency-specific cDNA from barley roots having two homologous cys
 A:Reference number: S17299; MUID:91355948; PMID:1832055
 A:Accession: S17299
 A:Molecule type: mRNA
 A:Residues: 1-74 <OKU>
 A:Cross-references: EMBL:X58540; NID:919006; PIDN:CAA41432.1; PID:919007
 C:Genetics:
 A:Gene: ids-1
 C:Superfamily: metallochionein
 C:Keywords: metal binding

Query Match
 Best Local Similarity 75.7%; Score 53; DB 1; Length 74;
 Matches 9; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 VILGVPKIOPEG 13
 |||||
 Db 37 VILGVPKIOPEG 49

RESULT 6

S37239
 metallochionein-like protein - white clover
 C:Species: Trifolium repens (white clover)
 C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 20-Aug-1999
 C:Accession: S37239
 R:Ellison, N.W.
 submitted to the EMBL Data Library, September 1993
 A:Description: Sequence analysis of two cDNA clones for metallochionein-like proteins
 A:Reference number: S37239
 A:Accession: S37239
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-77 <ELL>
 A:Cross-references: EMBL:226492; NID:9403326; PIDN:CAA81264.1; PID:9403327
 C:Superfamily: metallochionein

Query Match
 Best Local Similarity 75.7%; Score 53; DB 2; Length 77;
 Matches 9; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 VILGVPKIOPEG 14
 |||||
 Db 40 VILGVPKIOPEG 53

RESULT 7

JQ2128
 metallochionein - soybean
 C:Species: Glycine max (soybean)
 C:Date: 28-Aug-1985 #sequence_revision 07-Oct-1994 #text_change 23-Mar-1995
 C:Accession: JQ2128
 R:Kawashima, I.; Inokuchi, Y.; Chino, M.; Kimura, M.; Shimizu, N.
 Plant Cell Physiol. 32, 913-916, 1991
 A:Title: Isolation of a gene for a metallochionein-like protein from soybean.
 A:Reference number: JQ2128
 A:Accession: JQ2128
 A:Molecule type: mRNA
 A:Residues: 1-79 <KAW>
 A:Experimental source: seedling, cv. Saxa
 C:Comment: This protein participates in detoxification and metabolism of heavy metals
 C:Superfamily: metallochionein
 C:Keywords: metal binding

Query Match
 Best Local Similarity 70.0%; Score 49; DB 2; Length 79;

Best Local Similarity 57.1%; Pred. No. 0.1;
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 VIIIGVPAKIOFEG 14
: : : : :
Db 40 LIVGVAPKAKOFEFG 53

RESULT 8

T10087
metallothionein - castor bean
C:Species: Ricinus communis (castor bean)
C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 11-May-2000
C:Accession: T10087
R:Yangala, S.; Bailey-Serres, J.
A:Title: Nucleotide sequence of a maize (Zea mays L.) cDNA (Accession No. U29383) coding
A:Reference number: Z16941; MUID:96030280; PMID:7480354
A:Accession: T10087
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-80 <VAN>
A:Cross-references: EMBL:L02306; NID:g169712; PID:g169713
A:Experimental source: strain Carmelita; tissue-type cotyledon
C:Genetics:
A:Gene: MTI
C:Superfamily: metallothionein
C:Keywords: metal binding

Query Match 61.4%; Score 43; DB 2; Length 80;
Best Local Similarity 57.1%; Pred. No. 1.3;
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 VIIIGVPAKIOFEG 14
: : : : :
Db 40 LIVGVAPKAKOFEFG 53

RESULT 9

S48038
metallothionein-like protein - kiwi fruit
C:Species: Actinidia chinensis var. deliciosa (kiwi fruit)
C:Date: 26-Dec-1994 #sequence_revision 27-Feb-1997 #text_change 20-Aug-1999
C:Accession: S48038
R:Leogier, S.E.; Gardner, R.C.
A:Title: Cloning and characterization of five cDNAs for genes differentially expressed
A:Reference number: S48035; MUID:94355660; PMID:8075403
A:Accession: S48038
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-78 <LED>
A:Cross-references: EMBL:L27813; NID:9450244; PIDN:AAA53074.1; PID:9450245
C:Superfamily: metallothionein

Query Match 60.0%; Score 42; DB 2; Length 78;
Best Local Similarity 57.1%; Pred. No. 1.9;
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 VIIIGVPAKIOFEG 14
: : : : :
Db 40 LIVGVAPKAKOFEFG 53

RESULT 10

F70569
hypothetical protein RV3489 - Mycobacterium tuberculosis (strain H37RV)
C:Species: Mycobacterium tuberculosis
C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999
C:Accession: F70569
R:Coile, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltham, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.

Nature 393, 537-544, 1998
A:Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete geno

A:Reference number: A70500; MUID:98295987; PMID:9634230
A:Accession: F70569
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-54 <COL>
A:Cross-references: GB:Z95390; GB:AL123456; NID:93261766; PIDN:CAB08712.1; PID:631608
A:Experimental source: strain H37RV
C:Genetics:
A:Gene: RV3489

Query Match 57.1%; Score 40; DB 2; Length 54;
Best Local Similarity 60.0%; Pred. No. 2.9;
Matches 6; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 2 IIVGVPAKIO 11
: : : : :
Db 44 MGIKPAKLE 53

RESULT 11

T07076
metallothionein type II B - tomato
C:Species: Lycopersicon esculentum (tomato)
C:Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 20-Apr-2000
C:Accession: T07076
R:Whitelaw, C.A.; Lehuquet, J.A.; Thurman, D.A.; Tomsett, A.B.
A:Title: The isolation and characterization of type II metallothionein-like gen
A:Reference number: Z15900
A:Accession: T07076
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-82 <WHT>
A:Cross-references: EMBL:L77966; NID:q1449137; PIDN:AAB04675.1; PID:q1449138
A:Experimental source: strain Ailsa craig
C:Genetics:
A:Gene: MTB

A:Introns: 22/2
C:Superfamily: metallothionein
C:Keywords: metal binding

Query Match 57.1%; Score 40; DB 2; Length 82;
Best Local Similarity 58.3%; Pred. No. 4.5;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 VIIIGVPAKIOF 12
: : : : :
Db 41 LIVGVPEKTSF 52

RESULT 12

T07114
metallothionein-like protein - tomato
C:Species: Lycopersicon esculentum (tomato)
C:Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 21-Jul-2000
C:Accession: T07114
R:Grilich, A.; Ganai, M.; Stephan, U.W.; Baumelein, H.
A:Title: Structure, expression and chromosomal localization of the metallothionein-11
A:Reference number: Z15923; MUID:98349862; PMID:9687073
A:Accession: T07114
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-82 <GIR>

A:Cross-references: EMBL:Z68138; NID:q1103688; PIDN:CAA92243.1; PID:q1103689
A:Experimental source: cultivar Bonner Best; root
C:Superfamily: metallothionein
C:Keywords: metal binding

Query Match 57.1%; Score 40; DB 2; Length 82;

Best Local Similarity 58.3%; Pred. No. 4.5;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 LVIGVPAKIOF 12
:|||||
Db 41 LVIGVPEKTSF 52

RESULT 13

F91002

Probable prophage repressor CI [imported] - Escherichia coli (strain O157:H7, substrain C:Species: Escherichia coli

C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 03-Jun-2002

C:Accession: F91002

R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.; gasawara, N.; Yasunaga, T.; Khara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.

DNA Res. 8, 11-22, 2001

A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gene A:Reference number: A99629; MUID:21156231; PMID:11258796

A:Accession: F91002

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-212 <HAY>

A:Cross-references: GB:BA000007; PIDN:BA36413.1; PID:G13362459; GSPDB:GN00154

A:Experimental source: strain O157:H7, substrain RMD 0509952

C:Genetics:

A:Gene: ECS2990

C:Superfamily: repressor protein ci

Query Match

Best Local Similarity 80.0%; Pred. No. 12;

Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

RESULT 14

E85847

hypothetical protein 23358 [imported] - Escherichia coli (strain O157:H7, substrain EDL9

C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 03-Jun-2002

C:Accession: E85847

R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew

Iller, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,

Nature 409, 529-533, 2001

A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.

A:Reference number: A85480; MUID:21074935; PMID:11206551

A:Accession: E85847

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-215 <STO>

A:Cross-references: GB:AE005174; NID:912516411; PIDN:AG57241.1; GSPDB:GN00145; UWGP:233

A:Experimental source: strain O157:H7, substrain EDL933

C:Genetics:

A:Gene: 23358

C:Superfamily: repressor protein ci

Query Match

Best Local Similarity 80.0%; Pred. No. 12;

Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

RESULT 15

AE0164

Probable exported protein YPO1347 [imported] - Yersinia pestis (strain CO92)

C:Species: Yersinia pestis

C:Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 02-Nov-2001

C:Accession: AE0164
R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Tillett, R.W.; Holden, M.T.G.; Prentice, M.
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.
il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell
Nature 413, 523-527, 2001

A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.

A:Reference number: AB0001; MUID:21470413; PMID:11586360

A:Accession: AE0164

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-456 <KUR>

A:Cross-references: GB:AL590842; PIDN:CAC90176.1; PID:915979395; GSPDB:GN00175

C:Genetics:

A:Gene: YPO1347

Query Match

Best Local Similarity 72.7%; Pred. No. 26;

Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 LVIGVPAKIO 11
:|||||

Db 8 LILGAGPMAIO 18

Search completed: January 13, 2003, 09:57:06
Job time: 14.9231 secs

Mon Jan 13 10:35:09 2003

us-09-554-941-1.rsp

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 13, 2003, 09:53:15 : Search time 14 Seconds
(without alignments)
41.476 Million cell updates/sec

Title: us-09-554-941-1
Perfect score: 70
Sequence: 1 VILGVPKIOFEF 14

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	70	100.0	MT1B_VICFA	Q41670 vicia faba
2	70	100.0	MT1A_VICFA	P20830 pisum sativ
3	64	91.4	MTB_TRIIP	P43399 trifolium r
4	60	85.7	MT1A_VICFA	Q39458 cicor ariet
5	56	80.0	MT1A_VICFA	Q41669 vicia faba
6	53	75.7	MT1A_VICFA	P26571 hordeum vul
7	53	75.7	MT1A_VICFA	P43398 trifolium r
8	43	61.4	MT2_VICFA	P30564 ricinus com
9	42	60.0	MT2_VICFA	P43390 actinidia c
10	41	58.6	MT2_VICFA	Q39459 cicor ariet
11	41	58.6	AA1M_MOUSE	Q91K77 mus musculu
12	40	57.1	AA1M_MOUSE	Q00189 homo sapien
13	40	57.1	MT2B_LYCES	Q40158 lycopersico
14	39	55.7	VIB5_AGRF5	P17795 arabidateri
15	38	54.3	MT2_VICFA	Q41657 vicia faba
16	38	54.3	MT2_VICFA	P94029 oryza sativ
17	38	54.3	6PBD_PIG	P14332 sus scrofa
18	38	54.3	SAT1_SYNY3	P74241 synecocyst
19	38	54.3	YLS4_CABEL	P34389 caenorhabd
20	37	52.9	NUSG_AOUAR	O67757 aquilex aeo
21	37	52.9	MRS4_YEAST	P23500 saccharomyc
22	37	52.9	MRS3_YEAST	P10566 saccharomyc
23	37	52.9	FRZG_MYXXA	P31758 myxococcus
24	37	52.9	ADH1_APTAU	P46645 apteryx aus
25	37	52.9	PME_MEDSA	Q42920 medicago sa
26	37	52.9	LRG1_YEAST	P35688 saccharomyc
27	37	52.9	ATRS2_BOVIN	P79331 b adams-2
28	37	52.9	ATRS2_BOVIN	O95450 h adams-2
29	36	51.4	MCA2_DROME	O94X99 mus musculu
30	36	51.4	TIN2_MOUSE	P05455 homo sapien
31	36	51.4	LA_HUMAN	P52209 homo sapien
32	36	51.4	6PGD_HUMAN	P26204 trifolium r
33	36	51.4	BGIS_TRIIP	

34	36	51.4	803	1	Z151_HUMAN
35	35	50.0	78	1	MT2_MUSAC
36	35	50.0	344	1	COMC_METUA
37	35	50.0	366	1	PHD1_YEAST
38	35	50.0	374	1	PME_PETIN
39	35	50.0	443	1	DCUA_HELPJ
40	35	50.0	443	1	DCUA_HELPJ
41	35	50.0	458	1	OOAL_RHIME
42	35	50.0	463	1	STHA_PSEAE
43	35	50.0	529	1	GUAA_MYCLE
44	35	50.0	621	1	DCTB_RHIME
45	35	50.0	643	1	NOSZ_ALCEU

ALIGNMENTS

RESULT 1	ID	MT1B_VICFA	STANDARD	PRT	75 AA
AC	Q41670				
DT	01-NOV-1997 (Rel. 35, Created)				
DT	01-NOV-1997 (Rel. 35, Last sequence update)				
DT	15-JUN-2002 (Rel. 41, Last annotation update)				
DE	Metallothionein-like protein 1B.				
GN	MT1B.				
OS	Vicia faba (Broad bean).				
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;				
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; Rosidae;				
OC	euroside 1; Fabales; Fabaceae; Papilionoideae; Viciae; Vicia.				
OX	NCBI_TaxID=3906;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=leaf;				
RX	MEDLINE=97238465; PubMed=9132050;				
RA	Foley R.C., Liang Z.M., Singh K.B.;				
RT	"Analysis of type 1 metallothionein cDNAs in Vicia faba."				
RL	Plant Mol. Biol. 33:583-591(1997).				
CC	-1- FUNCTION: METALLOTHIONEIN HAVE A HIGH CONTENT OF CYSTEINE				
CC	-1- RESIDUES THAT BIND VARIOUS HEAVY METALS.				
CC	-1- SIMILARITY: BELONGS TO THE METALLOTHIONEIN SUPERFAMILY; FAMILY 15.				
CC	-1- This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/				
CC	or send an email to license@sib-sib.ch).				
DR	EMBL: X91078; CA62552.1;				
DR	InterPro: IPR000347; Metallothion_15.				
DR	Pfam: PF01439; Metallothio_2; 1.				
DR	ProDom: PD001611; Metallothion_15; 1.				
KW	Metal-binding; Metal-thiolate cluster; Multigene family.				
SQ	SEQUENCE 75 AA; 7750 MW; P59FC15A6A025BAA CRC64;				
Query Match	100.0%; Score 70; DB 1; Length 75;				
Best Local Similarity	100.0%; Pred. No. 1.2e-05;				
Matches 14; Conservative	0; Mismatches 0; Indels 0; Gaps 0;				
QY	1 VILGVPKIOFEF 14				
DB	36 VILGVPKIOFEF 49				
RESULT 2					
ID	MT1A_VICFA	STANDARD	PRT	75 AA	
AC	P20830				
DT	01-FEB-1991 (Rel. 17, Created)				
DT	01-FEB-1991 (Rel. 17, Last sequence update)				
DT	15-JUN-2002 (Rel. 41, Last annotation update)				

DE Metallothionein-like protein 1.
 GN MTA.
 OS Pisum sativum (Garden pea).
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 CC eurosids I; Fabales; Fabaceae; Papilionoideae; Viciae; Pisum.
 RN NCBI_TaxID=3888;
 RP [1]
 RC SEQUENCE FROM N.A.
 RC STRAIN=cv. Feltham First; TISSUE=Root;
 RA MEDLINE=90201367; PubMed=2318309;
 RA Evans I.M., Gatehouse L.N., Gatehouse J.A., Robinson N.J.,
 RA Croy R.R.D.;
 RT "A gene from pea (Pisum sativum L.) with homology to metallothionein
 RT genes";
 RL FEBS Lett. 262:29-32(1990).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=9211733; PubMed=1765150;
 RA Kille P., Winge D.R., Harwood J.L., Kay J.;
 RT "A plant metallothionein produced in E. coli.";
 RL FEBS Lett. 295:171-175(1991).
 CC -!- FUNCTION: METALLOTHIONEINS HAVE A HIGH CONTENT OF CYSTEINE
 CC RESIDUES THAT BIND VARIOUS HEAVY METALS.
 CC -!- SIMILARITY: BELONGS TO THE METALLOTHIONEIN SUPERFAMILY; FAMILY 15.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: Z23097; CAAB0645.1; -
 DR PIR: S09098; S09098.
 DR PIR: S20347; S20347.
 DR InterPro: IPR000347; Metallothion_15.
 DR Pfam: PF01439; Metallothio_2; 1.
 DR ProDom: PD001611; Metallothion_15; 1.
 KW Metal-binding; Metal-thiolate cluster.
 SQ SEQUENCE 75 AA; 7608 MW; 68856F6F6883A70 CRC64;
 Query Match Best Local Similarity 100.0%; Score 70; DB 1; Length 75;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 VILVGPAKIOFE 14
 DB 36 VILVGPAKIOFE 49
 RESULT 3
 MTB_TRIRP
 ID MTB_TRIRP STANDARD: PRT; 75 AA.
 AC P43399;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Metallothionein-like protein B.
 GN MTB.
 OS Trifolium repens (Creeping white clover).
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 CC eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae; Trifolium.
 RN NCBI_TaxID=3899;
 RP [1]
 RC SEQUENCE FROM N.A.
 RC STRAIN=cv. Huia; TISSUE=Stolon node;
 RA Ellison N.W., White D.W.R.;
 RT "Isolation of two cDNA clones encoding metallothionein-like proteins
 RT from Trifolium repens L.";
 RL (in) Plant Gene Register PGR96-068.

CC -!- FUNCTION: METALLOTHIONEINS HAVE A HIGH CONTENT OF CYSTEINE
 CC RESIDUES THAT BIND VARIOUS HEAVY METALS.
 CC -!- SIMILARITY: BELONGS TO THE METALLOTHIONEIN SUPERFAMILY; FAMILY 15.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: Z26493; CAAB1265.1; -
 DR InterPro: IPR000347; Metallothion_15.
 DR Pfam: PF01439; Metallothio_2; 1.
 DR ProDom: PD001611; Metallothion_15; 1.
 KW Metal-binding; Metal-thiolate cluster; Multigene family.
 SQ SEQUENCE 75 AA; 7659 MW; 64ACACCC234F54FE CRC64;
 Query Match Best Local Similarity 91.4%; Score 64; DB 1; Length 75;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 VILVGPAKIOFE 13
 DB 36 VILVGPAKIOFE 48
 RESULT 4
 MT1_CICAR
 ID MT1_CICAR STANDARD: PRT; 75 AA.
 AC Q39436;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Metallothionein-like protein 1 (MT-1).
 OS Cicer arietinum (Chickpea) (Garbanzo).
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 CC eurosids I; Fabales; Fabaceae; Papilionoideae; Ciceraceae; Cicer.
 RN NCBI_TaxID=3827;
 RP [1]
 RC SEQUENCE FROM N.A.
 RC TISSUE=Etolated epicotyl;
 RA Dopico B., Labrador E., Olan R.V., Munoz F.J.;
 RL Submitted (Apr-1996) to the EMBL/GenBank/DBJ databases
 CC -!- FUNCTION: METALLOTHIONEINS HAVE A HIGH CONTENT OF CYSTEINE
 CC RESIDUES THAT BIND VARIOUS HEAVY METALS.
 CC -!- SIMILARITY: BELONGS TO THE METALLOTHIONEIN SUPERFAMILY; FAMILY 15.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X95708; CAAB5008.1; -
 DR InterPro: IPR000347; Metallothion_15.
 DR Pfam: PF01439; Metallothio_2; 1.
 DR ProDom: PD001611; Metallothion_15; 1.
 KW Metal-binding; Metal-thiolate cluster.
 SQ SEQUENCE 75 AA; 7610 MW; 9806F0CB0B63F2BC CRC64;
 Query Match Best Local Similarity 85.7%; Score 60; DB 1; Length 75;
 Matches 11; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 QY 1 VILVGPAKIOFE 14
 DB 36 VILVGPAKIOFE 49

```

CC CC PHOSPHODIPHOSPHATE PRODUCTION IN FE-DEFICIENT BARLEY ROOTS.
CC CC MAY HAVE A FUNCTION AT THE REGULATORY REGION OF MAS SYNTHETIC
CC CC GENES OR FE(II)-MAS TRANSPORTER GENE BY CONJUGATING WITH FE(2+)
CC CC LIKE FOR PROTEIN.
CC CC -1- SIMILARITY: BELONGS TO THE METALLOTHIONEIN SUPERFAMILY; FAMILY 15.
CC CC -----
CC CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC CC or send an email to license@isb-sib.ch).
CC CC -----
CC CC EMBL: X58540; CAA1432.1; -
CC CC PIR: S15558; SMBH1.
CC CC InterPro: IPR000347; Metallothion_15.
CC CC Pfam: PF01439; Metallothio_2; 1.
CC CC ProDom: PD001611; Metallothion_15; 1.
CC CC Metal-binding; Metal-thiolate cluster.
CC CC SEQUENCE 74 AA; 7469 MW; F86C8C3F65901C6 CRC64;
CC CC -----
CC CC Query Match 75.7%; Score 53; DB 1; Length 74;
CC CC Best local similarity 69.2%; Pred. No. 0.012;
CC CC Matches 9; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
CC CC -----
QY 1 VIIIGVAPAKIOFE 13
DB 37 VIIIGVSAKIOFE 49
:::|||||:::
RESULT 7
MFA_TRIIP STANDARD; PRT; 77 AA.
ID MFA_TRIIP
AC P43398;
DC 01-NOV-1995 (Rel. 32, last sequence update)
DT 01-NOV-1995 (Rel. 32, last sequence update)
DI 15-JUN-2002 (Rel. 41, last annotation update)
DE Metallothionein-like protein A.
GN MT1A.
OS Trifolium repens (Creeping white clover).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids 1; Fabales; Fabaceae; Papilionoideae; Trifolieae; Trifolium.
OC NCBI_TaxId=3899;
OX [1]
RP SEQUENCE FROM N.A.
RN SPRAIN-CV. Huia; TISSUE=Stolon node;
RA Ellison N.W., White D.W.R.;
RT Isolation of two cDNA clones encoding metallothionein-like proteins
RT from Trifolium repens L.;
RL (in) Plant Gene Register PGR96-068.
RT FUNCTION: METALLOTHIONEINS HAVE A HIGH CONTENT OF CYSTEINE
-1- RESIDUES THAT BIND VARIOUS HEAVY METALS.
CC -1- SIMILARITY: BELONGS TO THE METALLOTHIONEIN SUPERFAMILY; FAMILY 15.
CC CC -----
CC CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC CC or send an email to license@isb-sib.ch).
CC CC -----
CC CC EMBL: Z26492; CAA81264.1; -
CC CC InterPro: IPR000347; Metallothion_15.
CC CC Pfam: PF01439; Metallothio_2; 1.
CC CC ProDom: PD001611; Metallothion_15; 1.
CC CC Metal-binding; Metal-thiolate cluster; Multigene family.
CC CC SEQUENCE 77 AA; 7503 MW; 5FAC163BA644B8D1 CRC64;
CC CC -----
CC CC Query Match 75.7%; Score 53; DB 1; Length 77;
CC CC Best local similarity 64.3%; Pred. No. 0.012;

```

Matches 9; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
 QY 1 VILGVPKIOPEG 14
 Db 40 LVIGVGAERARPEG 53

RESULT 8

MT2_RICCO STANDARD: PRT; 80 AA.

AC P30364;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DE Metallothionein-like protein type 2.
 GN MT1.
 OS Ricinus communis (Castor bean).
 CC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
 CC Eumetazoa; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 CC Eucosids I; Malpighiales; Euphorbiaceae; Ricinus.
 OX NCBI_TaxID=3988;
 RN (1)
 RP SEQUENCE FROM N.A.
 RA Weig A.; Komor E.;
 RT "Isolation of a class II metallothionein cDNA from Ricinus communis
 L.";
 RL (1) Plant Gene Register PGR95-066.
 CC -1- FUNCTION: METALLOTHIONEIN HAVE A HIGH CONTENT OF CYSTEINE
 CC -1- RESIDUES THAT BIND VARIOUS HEAVY METALS.
 CC -1- SIMILARITY: BELONGS TO THE METALLOTHIONEIN SUPERFAMILY: FAMILY 15.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; L02306; AAC37473.1;
 DR InterPro; IPR000347; Metallothion_15.
 DR Pfam; PF01439; Metallothio_2; 1.
 DR ProDom; PD001611; Metallothion_15; 1.
 DR KW Metal-binding; Metal-thiolate cluster.
 SQ SEQUENCE 80 AA; 7953 MW; AA7304254491A3B7 CRC64;

Query Match
 Best Local Similarity 57.1%; Score 43; DB 1; Length 80;
 Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
 QY 1 VILGVPKIOPEG 14
 Db 40 LVIGVGAERARPEG 53

RESULT 9

MT2_ACTCH STANDARD: PRT; 78 AA.

AC P43390;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DE Metallothionein-like protein type 2 (Kiwifruit).
 GN Actinidia chinensis (Kiwifruit) (Yangtze).
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Asteridae; Ericales; Actinidiaceae; Actinidia.
 OX NCBI_TaxID=3625;
 RN (1)
 RP SEQUENCE FROM N.A.
 RA STRAIN-ev. Deliclosa Hayward; TISSUP=fruit;
 RX MEDLINE=9435560; PubMed=8075403;
 RA Ledger S.E.; Gardner R.C.;

RT "Cloning and characterization of five cDNAs for genes differentially
 RT expressed during fruit development of kiwifruit (Actinidia deliciosa
 var. deliciosa).";
 RL Plant Mol. Biol. 25:877-886(1994).
 CC -1- FUNCTION: METALLOTHIONEIN HAVE A HIGH CONTENT OF CYSTEINE
 CC -1- RESIDUES THAT BIND VARIOUS HEAVY METALS.
 CC -1- DEVELOPMENTAL STAGE: HIGHLY EXPRESSED IN YOUNG FRUIT WITH REDUCED
 CC EXPRESSION IN THE LATER STAGES OF FRUIT DEVELOPMENT.
 CC -1- SIMILARITY: BELONGS TO THE METALLOTHIONEIN SUPERFAMILY: FAMILY 15.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; L27813; AAA53074.1;
 DR InterPro; IPR000347; Metallothion_15.
 DR Pfam; PF01439; Metallothio_2; 1.
 DR ProDom; PD001611; Metallothion_15; 1.
 DR KW Metal-binding; Metal-thiolate cluster.
 SQ SEQUENCE 78 AA; 7828 MW; E2FE9BE0FFD501D CRC64;

Query Match
 Best Local Similarity 50.0%; Score 42; DB 1; Length 78;
 Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
 QY 1 VILGVPKIOPEG 14
 Db 40 LVIGVGAERARPEG 53

RESULT 10

MT2_CICAR STANDARD: PRT; 79 AA.

AC Q39459;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Metallothionein-like protein 2 (WF-2).
 GN Cicer arietinum (Chickpea) (Garbanzo).
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Eumetazoa; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 CC Eucosids I; Fabales; Fabaceae; Papilionoideae; Ciceraceae; Cicer.
 OX NCBI_TaxID=3827;
 RN (1)
 RP SEQUENCE FROM N.A.
 RA TISSUE=Etliolated epicotyl;
 RA Dopico B.; Labrador E.; Gillan R.V.; Munoz F.J.;
 RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: METALLOTHIONEIN HAVE A HIGH CONTENT OF CYSTEINE
 CC -1- RESIDUES THAT BIND VARIOUS HEAVY METALS.
 CC -1- SIMILARITY: BELONGS TO THE METALLOTHIONEIN SUPERFAMILY: FAMILY 15.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; X57509; CA65009.1;
 DR InterPro; IPR000347; Metallothion_15.
 DR Pfam; PF01439; Metallothio_2; 1.
 DR ProDom; PD001611; Metallothion_15; 1.
 DR KW Metal-binding; Metal-thiolate cluster; Multigene family.
 SQ SEQUENCE 79 AA; 7947 MW; 0DBE6C8DBF356071 CRC64;

Query Match
 Best Local Similarity 58.6%; Score 41; DB 1; Length 79;
 50.0%; Pred. No. 1.7;

Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

OY 1 VILGVPKIOFEG 14
 DB 40 LVMGVASKTOFEG 53

RESULT 11

A4M1_MOUSE STANDARD; PRT; 449 AA.

AC 09JCK7; 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Adapter-related protein complex 4 mu 1 subunit (Mu subunit of Ap-4)
 DE (Ap-4 adapter complex mu subunit) (Mu-adaptin-related protein 2) (mu-
 DE ARP2) (mu4).
 GN AP4M1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;

RP SEQUENCE FROM N.A.

RA Werner H., Nave K., A.;
 RT "Mus musculus adaptor-related protein complex Ap-4 mu4 subunit.";
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.

CC -1- FUNCTION: SUBUNIT OF NOVEL TYPE OF CLATHRIN-OR NON-CLATHRIN-
 CC ASSOCIATED PROTEIN COAT INVOLVED IN TARGETING PROTEINS FROM THE
 CC TRANS-GOLGI NETWORK (TGN) TO THE ENDOSOMAL-LYSOSOMAL SYSTEM (BY
 CC SIMILARITY).
 CC -1- SUBUNIT: ADAPTER-LIKE COMPLEX 4 (AP-4) IS AN HETEROETRAMER
 CC COMPOSED OF TWO LARGE CHAINS (EPSILON/AP4E1 AND BETA/AP4B1), A
 CC MEDIUM CHAIN (MU/AP4M1) AND A SMALL CHAIN (SIGMA/AP4S1) (BY
 CC SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: ASSOCIATED WITH THE TRANS-GOLGI NETWORK (BY
 CC SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE ADAPTOR COMPLEXES MEDIUM SUBUNITS
 CC FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL: AF242858; AAF63513.1;
 DR MGD: MGI:1337063; AP4M1.
 DR InterPro: IPR001392; Clathrn_med.
 DR Pfam: PF00928; Adap_comp_sub; 1.
 DR PRINTS: PR00314; CLATHRNADPT.
 DR PROSITE: PS00990; CLAT_ADAPTOR_M.1; FALSE_NEG.
 DR PROSITE: PS00991; CLAT_ADAPTOR_M.2; FALSE_NEG.
 KM Coated pits; Endocytosis.
 SQ SEQUENCE 449 AA; 49509 MW; 579A9E1F8255D04 CRC64;

Query Match 58.6%; Score 41; DB 1; Length 449;
 Best Local Similarity 63.6%; Pred. No. 9;
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 3 LGVGPKIOFEG 13
 DB 396 LGVGPASLSFE 406

RESULT 12

A4M1_HUMAN STANDARD; PRT; 453 AA.

ID A4M1_HUMAN
 AC 000189; 09OHR9;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)

DE 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Adapter-related protein complex 4 mu 1 subunit (Mu subunit of Ap-4)
 DE (Ap-4 adapter complex mu subunit) (Mu-adaptin-related protein 2) (mu-
 DE ARP2) (mu4).
 GN AP4M1 OR MURP2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 OX NCBI_TaxID=9606;

RP SEQUENCE FROM N.A.

RC TISSUE-Brain; Pubmed-9013859;
 RA Wang X., Kilmann M.W.;
 RT "Identification of a fourth adaptor-related proteins, mu-ARP1 and
 RT mu-ARP2.";
 RL FEBS Lett. 402:57-61(1997).

RP SEQUENCE FROM N.A.

RC MEDLINE-9936516; Pubmed-10436028;
 RA Hirst J., Bright N.A., Rous B., Robinson M.S.;
 RT "Characterization of a fourth adaptor-related protein complex.";
 RL Mol. Biol. Cell 10:2787-2802(1999).

RP SEQUENCE FROM N.A., AND CHARACTERIZATION.

RC TISSUE-Spleen;
 RA Aquilar R.C., Boehm M., Gorskova I., Crouch R.J., Tomita K.,
 RT Salto T., Ono H., Bonifacio J.S.;
 RL "Signal-binding specificity of the mu4 subunit of the adaptor protein
 RL complex, Ap-4.";
 RL J. Biol. Chem. 276:13145-13152(2001).

CC -1- FUNCTION: SUBUNIT OF NOVEL TYPE OF CLATHRIN-OR NON-CLATHRIN-
 CC ASSOCIATED PROTEIN COAT INVOLVED IN TARGETING PROTEINS FROM THE
 CC TRANS-GOLGI NETWORK (TGN) TO THE ENDOSOMAL-LYSOSOMAL SYSTEM.
 CC -1- SUBUNIT: ADAPTER-LIKE COMPLEX 4 (AP-4) IS AN HETEROETRAMER
 CC COMPOSED OF TWO LARGE CHAINS (EPSILON/AP4E1 AND BETA/AP4B1), A
 CC MEDIUM CHAIN (MU/AP4M1) AND A SMALL CHAIN (SIGMA/AP4S1)
 CC -1- SUBCELLULAR LOCATION: ASSOCIATED WITH THE TRANS-GOLGI NETWORK.
 CC -1- TISSUE SPECIFICITY: UBIQUITOUS. HIGHLY EXPRESSED IN TESTIS AND
 CC LOWLY EXPRESSED IN BRAIN AND LUNG.

CC -1- DOMAIN: INTERACTS SPECIFICALLY WITH TYROSINE-BASED SORTING
 CC SIGNALS.
 CC -1- SIMILARITY: BELONGS TO THE ADAPTOR COMPLEXES MEDIUM SUBUNITS
 CC FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL: Y08387; CAA69667.1;
 DR EMBL: AF15158; AAD4328.1;
 DR EMBL: AF020796; AAD2589.1;
 DR Genew: HGNC:574; AP4M1.
 DR MIM: 602296;
 DR InterPro: IPR001392; Clathrn_med.
 DR Pfam: PF00928; Adap_comp_sub; 1.
 DR PROSITE: PS00990; CLAT_ADAPTOR_M.1; FALSE_NEG.
 DR PROSITE: PS00991; CLAT_ADAPTOR_M.2; FALSE_NEG.
 DR Coated pits; Endocytosis.
 KW Coated pits; 338
 FT CONFLICT 338 R -> G (IN REF. 2).
 FT CONFLICT 345 L -> Q (IN REF. 2).
 FT CONFLICT 400 L -> M (IN REF. 2).
 FT CONFLICT 417 S -> C (IN REF. 2).
 SQ SEQUENCE 453 AA; 50005 MW; 138B13A8C91D3444 CRC64;

Query Match 58.6%; Score 41; DB 1; Length 453;
 Best Local Similarity 63.6%; Pred. No. 9.1;

Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 QY 3 LGVGPARIQF 13
 DB 400 LGVGPASISFE 410

RESULT 13
 ID MT2B.LYCES STANDARD; PRT; 82 AA.
 AC 040158; 043514;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Metallothionein-like protein type 2 B.
 GN MTB.
 OS Lycopersicon esculentum (Tomato).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
 OX NCBI_TaxID:4081;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CV. Ailsa Craig;
 RA Whitehead C.A., Lehuquet J.A., Thurman D.A., Tomsett A.B.;
 RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
 RP SEQUENCE FROM N.A.
 RC STRAIN-CV. Bonnet Beste; TISSUE=Root;
 RA Girlich A., Herdik A., Balzer H., Stephan U., Baumlein H.;
 RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: METALLOTHIONEIN HAVE A HIGH CONTENT OF CYSTEINE
 CC RESIDUES THAT BIND VARIOUS HEAVY METALS.
 CC -1- SIMILARITY: BELONGS TO THE METALLOTHIONEIN SUPERFAMILY; FAMILY 15.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 DR EMBL; L77966; AAB04675.1;
 DR EMBL; Z68138; CAA92243.1;
 DR InterPro; IPR000347; Metallothion_15.
 DR Pfam; PF01439; Metallothio_2; 1.
 DR Prodom; PD001611; Metallothion_15; 1.
 KW Metal-binding; Metal-thiolate cluster; Multigene family.
 FT CONFLICT 17 17 G->D (IN REF. 2).
 SQ SEQUENCE 82 AA; 8253 MW; 6B298C4915CE495F CRC64;
 Query Match 57.1%; Score 40; DB 1; Length 82;
 Best Local Similarity 58.3%; Pred. No. 2.6;
 Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
 QY 1 VILGVPARIQF 12
 DB 41 LVGVGPERSISFE 52

OC Rhizobiaceae; Rhizobium.
 OX NCBI_TaxID=176299;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=90318324; PubMed=2370849;
 RA Kuldau G.A., de Vos G., Owen J., McAffrey G., Zambryski P.;
 RT "The virB operon of Agrobacterium tumefaciens pTiC58 encodes 11 open
 RT reading frames";
 RL Mol. Gen. Genet. 221:256-266(1990).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=90301800; PubMed=2194232;
 RA Rogowsky P.M., Powell B.S., Shirsar K., Lin T.-S., Morel P.,
 RA Zyprian E.M., Steck T.R., Kado C.I.;
 RT "Molecular characterization of the vir regulon of Agrobacterium
 RT tumefaciens: complete nucleotide sequence and gene organization of
 RL the 28.63-kbp regulon cloned as a single unit";
 RN [3]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=21608550; PubMed=11743193;
 RA Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,
 RA Okura Y.K., Zhou Y., Chen L., Wood G.E., Almeida N.F. Jr., Woo L.,
 RA Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Boyce D. Jr.,
 RA Chapman P., Clendenning J., Deatherage G., Gillet W., Grant C.,
 RA Kutayavin T., Levy R., Li M.-J., McClelland E., Palmeri A.,
 RA Raymond C., Rouse G., Saenphimmachak C., Wu Z., Romero P., Gordon D.,
 RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,
 RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
 RA Chumley F., Tingey S.V., Tomb J.-F., Gordon K.P., Olson M.V.,
 RA Westler E.W.;
 RT "The genome of the natural genetic engineer Agrobacterium tumefaciens
 RT C58";
 RL Science 294:2317-2323(2001).
 RN [4]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=21608551; PubMed=11743194;
 RA Goodner B., Hunkle G., Gattung S., Miller N., Blanchard M.,
 RA Quorillo B., Goldman B.S., Cao Y., Askenazi M., Halling C., Mullin L.,
 RA Houmel K., Gordon J., Vaudin M., Tarchouk O., Epp A., Liu F.,
 RA Woliam C., Allinger M., Dougherty D., Scott C., Lappas C., Markelz B.,
 RA Flanagan C., Crowell C., Gursen J., Lomo C., Sear C., Strub G.,
 RA Cleo C., Slater S.;
 RT "Genome sequence of the plant pathogen and biotechnology agent
 RT Agrobacterium tumefaciens C58";
 RL Science 294:2323-2328(2001).
 CC -1- FUNCTION: VIRB PROTEINS ARE SUGGESTED TO ACT AT THE BACTERIAL
 CC SURFACE AND THERE PLAY AN IMPORTANT ROLE IN DIRECTING T-DNA
 CC TRANSFER TO PLANT CELLS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 DR EMBL; X53264; CAA37358.1;
 DR EMBL; J03320; AAB91595.1;
 DR EMBL; AE009435; AAL46407.1;
 DR EMBL; AE007923; AAK90933.1;
 KW PIR; S12345; B5AG58.
 FT CROWN gall tumor; Plasmid; Signal; Complete proteome.
 FT SIGNAL 1 23 POTENTIAL.
 FT CHAIN 24 220 VIRB5 PROTEIN.
 SQ SEQUENCE 220 AA; 23269 MW; 78FEF913E03EA207 CRC64;
 Query Match 55.7%; Score 39; DB 1; Length 220;
 Best Local Similarity 63.6%; Pred. No. 10;
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 QY 2 ILGVGPARIQF 12

Db 15 LISTGPAPARQF 25

!!!!!!!!

```

RESULT 15
MT2_VICFA          STANDARD;          PRT;          77 AA.
ID MT2_VICFA
AC 041657;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Metallothionein-like protein type 2.
GN MT1.
OS Vicia faba (Broad bean).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Viciae; Vicia.
OC NCBI_TaxID=3906;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95036014; PubMed=7948889;
RA Foley R.C., Singh K.B.;
RT "Isolation of a Vicia faba metallothionein-like gene: expression in
RT foliar trichomes."
RT Plant Mol. Biol. 26:435-444(1994).
CC -I- FUNCTION: METALLOTHIONEIN HAVE A HIGH CONTENT OF CYSTEINE
CC RESIDUES THAT BIND VARIOUS HEAVY METALS.
CC -I- TISSUE SPECIFICITY: EXPRESSED IN THE LEFT, STEM AND FLOWER, AT
CC VERY LOW LEVELS IN ROOTS AND IS NOT DETECTABLE IN MESOPHYLL
CC PROTOPLASTS.
CC -I- SIMILARITY: BELONGS TO THE METALLOTHIONEIN SUPERFAMILY; FAMILY 15.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
-----
CC EMBL: X77254; CAA54471.1;
CC InterPro: IPR000347; Metallothion_15.
CC Pfam: PF01439; Metallothio_2; 1.
CC DR ProDom: PD001611; Metallothion_15; 1.
CC KW Metal-binding; Metal-thiolate cluster; Multigene family.
CC SEQUENCE 77 AA; 7730 MW; 7D7BAD8D023BF3C CRC64;
SO

Query Match 54.38; Score 38; DB 1; Length 77;
Best Local Similarity 53.88; Pred. No. 5-6;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
QY 1 VILGVPKIQFE 13
:||||| | | |
DB 40 LIMGVSEKAYE 52

```

Search completed: January 13, 2003, 09:54:05
Job time : 16.3333 secs

Mon Jan 13 10:35:10 2003

us-09-554-941-11.faq

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model
January 13, 2003, 09:53:15 : Search time 31.9487 Seconds
(without alignments)
58.391 Million cell updates/sec

Run on: US-09-554-941-11
Title: 73
Sequence: 1 OASIEGRTAPQPO 14
Scoring table: BL0SUM62
Gapop 10.0, Gapext 0.5

Searched: 908470 seqs, 133250620 residues
Total number of hits satisfying chosen parameters: 908470
Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: /SID2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
2: /SID2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
3: /SID2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
4: /SID2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
5: /SID2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
6: /SID2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
7: /SID2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
8: /SID2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
9: /SID2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10: /SID2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
11: /SID2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12: /SID2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13: /SID2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14: /SID2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15: /SID2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16: /SID2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17: /SID2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18: /SID2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19: /SID2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20: /SID2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
21: /SID2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22: /SID2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
23: /SID2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*

pred. No. is the number of results predicted by chance to have a
score greater than or equal to the total score distribution.
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	73	100.0	14	AAV22920	Immunoglobulin 11n
2	45	61.6	77	AA668281	Human vlgare ab
3	43	58.9	283	ABF03100	Human ORF protein
4	43	58.9	283	ABF03100	Drosophila melanog
5	42	57.5	105	ABR12935	Human secreted pro
6	42	57.5	313	ABR12935	Arabidopsis thalia
7	42	57.5	474	AA623805	Arabidopsis thalia
8	42	57.5	493	AA623805	Human polypeptide
9	41.5	56.8	49	AAO13603	Human polypeptide
10	41	56.2	22	ABR39729	Peptide #7235 enco

11	41	56.2	49	22	AAW73093
12	41	56.2	49	22	AAW73093
13	41	56.2	49	22	AAW73093
14	41	56.2	49	22	AAW73093
15	40.5	54.8	452	22	ABP26363
16	40	54.8	738	22	ABP44724
17	40	54.8	71	22	ABP43081
18	39	53.4	155	22	ABG05983
19	39	53.4	174	22	ABG61286
20	39	53.4	179	22	ABG05982
21	39	53.4	186	22	ABG94665
22	39	53.4	230	22	AAH47568
23	39	53.4	238	22	AAH83236
24	39	53.4	295	22	ABG52568
25	39	53.4	298	22	ABG55984
26	39	53.4	387	22	AAH06839
27	39	53.4	408	22	ABP30490
28	39	53.4	420	22	AAH22499
29	39	53.4	423	20	AAH23165
30	39	53.4	423	21	AAH39068
31	39	53.4	423	22	AAH39068
32	39	53.4	423	22	AAH39068
33	39	53.4	423	22	AAH39068
34	39	53.4	423	22	AAH39068
35	39	53.4	423	22	AAH39068
36	39	53.4	423	22	AAH39068
37	39	53.4	423	22	AAH39068
38	39	53.4	423	22	AAH39068
39	39	53.4	423	22	AAH39068
40	39	53.4	423	22	AAH39068
41	39	53.4	423	22	AAH39068
42	39	53.4	423	22	AAH39068
43	39	53.4	423	22	AAH39068
44	39	53.4	423	22	AAH39068
45	39	53.4	423	22	AAH39068

ALIGNMENTS

RESULT 1
AAV22920 standard; peptide: 14 AA.
ID AAV22920:
AC AAV22920:
XX 19-AUG-1999 (first entry)
XX Immunoglobulin linker used to make proteinase inhibitor fusions.
XX peptide linker: fusion protein: pathogen resistance, pathogen tolerance;
XX plant: transgene: proteinase inhibitor; Arabidopsis thaliana;
XX parasitic nematode.
XX Synthetic.
XX WO9928484-A1.
XX 10-JUN-1999.
XX 01-DEC-1998: 98NO-EP07792.
XX 03-DEC-1997: 97GB-0025556.
XX (NOVS) NOVARTIS AG.
XX Atkinson HJ, McPherson MJ, Urwin PE;
XX WPI: 1999-385387/32.
XX Proteinase inhibitor fusion proteins
XX Claim 11: Page 23: 39pp: English.
PS

XX The present sequence represents a peptide linker used in the fusion
 CC proteins of the invention. The specification describes a method for
 CC improving pathogen resistance or tolerance of a plant. The method
 CC comprises transfecting resistance or tolerance genes into a plant
 CC fusion protein which comprises two or more protein or domains
 CC expressed on their own. Specifically, two distinct proteins that
 CC inhibitors are co-delivered, at least one of which is a protein
 CC used to improve pathogen resistance or tolerance of a plant. The method is
 CC descendants, especially against parasitic nematode attack.

Sequence 14 AA:
 SO Query Match
 Best Local Similarity 100.0%; Score 73; DB 20; Length 14;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 QASIEGRYAPQPO 14
 DB 1 QASIEGRYAPQPO 14

RESULT 2

ID AAG68281
 AC AAG68281 standard; Protein: 321 AA.
 XX
 XX 18-FEB-2002 (first entry)
 DT
 DE Hordeum vulgare ab45 protein sequence SEQ ID NO:4.
 KW Hordeum vulgare cv. Himalaya; ab45; promoter; barley; plant;
 KM transcription; hydrolytic enzyme.
 OS Hordeum vulgare.
 PN US630656-B1.
 XX
 XX 23-OCT-2001.
 PD
 XX
 XX 13-OCT-1999; 99US-0417286.
 PF
 XX
 XX 13-OCT-1999; 99US-0417286.
 PR
 XX
 XX (SINT-) ACAD SINICA.
 PA
 XX
 XX Liu J, Cheng K, Chen T;
 PI WPI: 2002-040203/05.
 DR N-ESDB; ABA03926.
 XX
 XX

PT New barley promoter sequence useful for expressing heterologous
 PT proteins in plant cells, particularly hydrolytic enzymes in rice bran
 PT tissues to provide a cost-effective means of decreasing environmental
 PT impact of animal farming - Supplemental means of decreasing environmental
 PS Disclosure: Fig 1; 9pp; English.

CC The present invention describes an isolated barley ab45 promoter (1)
 CC comprising a 649 (ABA03924) or 587 (ABA03925) base pair sequence (1)
 CC bases 1 to 649 and bases 1 to 587 of ABA03926). Also described are:
 CC (1) a vector comprising (1); (2) an isolated nucleic acid comprising a
 CC the sequence that hybridizes to (ABA03925) under stringent conditions (also
 CC cells comprising the nucleic acid; and a plant cell; (3) transformed
 CC in plant cells, particularly in expressing heterologous proteins
 CC brain tissues, where expression of hydrolytic enzymes in rice
 CC may provide a cost-effective means of decreasing the environmental
 CC impact of animal farming. Supplemental means of decreasing environmental

CC farm animal feed can improve agricultural efficiency by decreasing the
 CC amount of fecal material, and subsequently the waste removal cost and
 CC environmental hazard. The present sequence represents the ab45 protein
 CC sequence given in the exemplification of the present invention.

Sequence 321 AA:
 SO Query Match
 Best Local Similarity 61.6%; Score 45; DB 23; Length 321;
 Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 QY 4 IEGRYTAPQ 13
 DB 75 LEGRYTAPQ 84

RESULT 3

ID ABP03100
 AC ABP03100 standard; Protein: 77 AA.
 XX
 XX 24-JUN-2002 (first entry)
 DT
 DE Human OREFX protein sequence SEQ ID NO:6182.
 KW Human; open reading frame; OREFX; gene therapy; cancer; cirrhosis;
 KM degenerative disorder; osteoarthritis; neurodegenerative disorder;
 KW cardiovascular disease; diabetes mellitus; systemic lupus erythematosus;
 KM immune deficiency; rheumatoid arthritis; autoimmune thyroiditis;
 KW myasthenia gravis.
 XX
 XX Homo sapiens.
 OS
 XX
 XX WO200192523-A2.
 PN
 XX
 XX 06-DEC-2001.
 PD
 XX
 XX 29-MAY-2001; 2001WO-US10836.
 PF
 XX
 XX 30-MAY-2000; 2000US-206132P.
 PR
 XX
 XX 29-AUG-2000; 2000US-228716P.
 PA
 XX
 XX (CURA-) CURAGEN CORP.
 PI
 XX
 XX Shinketsu RA, Leach MD;
 DR WPI: 2002-106308/14.
 XX N-PSDB; ABN18852.
 PT
 PT
 PT

PT Novel human polypeptides and polynucleotides useful for diagnosing,
 PT preventing and treating cardiovascular disease, neurodegenerative,
 PT hyperproliferative disorders and autoimmune disorders -
 PS Disclosure: SEQ ID 6182; 1037pp; English.

CC The present invention describes substantially purified human proteins
 CC (referred to as open reading frame OREFX, where X is 1-11491 (see Table 1
 CC in the specification). ABN15762 to ABN27252 encode the human OREFX
 CC proteins given in ABP00010 to ABP11500. OREFX proteins are useful for
 CC disorder or preventing a pathology associated with a human OREFX
 CC syndrome associated with OREFX-associated disorder. OREFX proteins are useful for
 CC treatment of cancer, hyperproliferative disorder, OREFX polynucleotide
 CC osteoarthritis, neurodegenerative disorders, cirrhosis of liver,
 CC transplanted, cardiovascular diseases, diabetes mellitus, systemic
 CC lupus erythematosus, hypertension, hypothyroidism, cholesterol ester
 CC storage disease, various immune deficiencies and disorders, infectious

PR 22-JUL-1999; 99US-0145085.
 PR 22-JUL-1999; 99US-0145087.
 PR 22-JUL-1999; 99US-0145089.
 PR 22-JUL-1999; 99US-0145192.
 PR 23-JUL-1999; 99US-0145218.
 PR 23-JUL-1999; 99US-0145226.
 PR 26-JUL-1999; 99US-0145913.
 PR 27-JUL-1999; 99US-0145918.
 PR 27-JUL-1999; 99US-0145919.
 PR 28-JUL-1999; 99US-0146386.
 PR 02-AUG-1999; 99US-0146388.
 PR 02-AUG-1999; 99US-0146389.
 PR 03-AUG-1999; 99US-0147038.
 PR 04-AUG-1999; 99US-0147302.
 PR 05-AUG-1999; 99US-0147192.
 PR 05-AUG-1999; 99US-0147260.
 PR 06-AUG-1999; 99US-0147303.
 PR 06-AUG-1999; 99US-0147416.
 PR 09-AUG-1999; 99US-0147493.
 PR 09-AUG-1999; 99US-0147935.
 PR 10-AUG-1999; 99US-0148171.
 PR 11-AUG-1999; 99US-0148319.
 PR 12-AUG-1999; 99US-0148341.
 PR 13-AUG-1999; 99US-0148565.
 PR 13-AUG-1999; 99US-0148684.
 PR 16-AUG-1999; 99US-0149368.
 PR 17-AUG-1999; 99US-0149175.
 PR 18-AUG-1999; 99US-0149426.
 PR 20-AUG-1999; 99US-0149722.
 PR 20-AUG-1999; 99US-0149723.
 PR 20-AUG-1999; 99US-0149929.
 PR 23-AUG-1999; 99US-0149902.
 PR 23-AUG-1999; 99US-0149930.
 PR 25-AUG-1999; 99US-0150566.
 PR 26-AUG-1999; 99US-0150884.
 PR 27-AUG-1999; 99US-0151065.
 PR 27-AUG-1999; 99US-0151066.
 PR 27-AUG-1999; 99US-0151080.
 PR 30-AUG-1999; 99US-0151303.
 PR 31-AUG-1999; 99US-0151438.
 PR 01-SEP-1999; 99US-0151930.
 PR 07-SEP-1999; 99US-0152363.
 PR 10-SEP-1999; 99US-0153070.
 PR 13-SEP-1999; 99US-0153758.
 PR 15-SEP-1999; 99US-0154018.
 PR 16-SEP-1999; 99US-0154039.
 PR 20-SEP-1999; 99US-0154779.
 PR 22-SEP-1999; 99US-0155139.
 PR 23-SEP-1999; 99US-0155486.
 PR 24-SEP-1999; 99US-0155659.
 PR 28-SEP-1999; 99US-0156456.
 PR 29-SEP-1999; 99US-0156596.
 PR 04-OCT-1999; 99US-0157117.
 PR 05-OCT-1999; 99US-0157753.
 PR 06-OCT-1999; 99US-0157865.
 PR 07-OCT-1999; 99US-0158029.
 PR 08-OCT-1999; 99US-0158232.
 PR 12-OCT-1999; 99US-0158369.
 PR 13-OCT-1999; 99US-0158993.
 PR 13-OCT-1999; 99US-0159284.
 PR 13-OCT-1999; 99US-0159295.
 PR 14-OCT-1999; 99US-0159329.
 PR 14-OCT-1999; 99US-0159330.
 PR 14-OCT-1999; 99US-0159637.
 PR 14-OCT-1999; 99US-0159638.
 PR 18-OCT-1999; 99US-0159584.
 PR 21-OCT-1999; 99US-0160741.
 PR 21-OCT-1999; 99US-0160767.

PR 21-OCT-1999; 99US-0160768.
 PR 21-OCT-1999; 99US-0160770.
 PR 21-OCT-1999; 99US-0160814.
 PR 21-OCT-1999; 99US-0160815.
 PR 22-OCT-1999; 99US-0160980.
 PR 22-OCT-1999; 99US-0160981.
 PR 22-OCT-1999; 99US-0160989.
 PR 25-OCT-1999; 99US-0161404.
 PR 25-OCT-1999; 99US-0161405.
 PR 25-OCT-1999; 99US-0161406.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161992.
 PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.

Query Match 57.5%; Score 42; DB 21; Length 313;
 Best local Similarity 53.8%; Pred. No. 25;
 Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

RESULT 7
 ID AAG3807 standard; Protein; 474 AA.

AC AAG3807;
 DT 17-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 27250.
 KW Protein identification: signal transduction pathway; metabolic pathway;
 KW hybridisation assay; genetic mapping; gene expression control; promoter;
 KW termination sequence.

OS Arabidopsis thaliana.
 PN EP1033405-A2.
 XX 06-SEP-2000.
 PD 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.
 PR 05-MAR-1999; 99US-0123180.
 PR 09-MAR-1999; 99US-0123548.
 PR 23-MAR-1999; 99US-0125788.
 PR 25-MAR-1999; 99US-0126264.
 PR 29-MAR-1999; 99US-0126785.
 PR 01-APR-1999; 99US-0127462.
 PR 06-APR-1999; 99US-0128234.
 PR 08-APR-1999; 99US-0128714.
 PR 16-APR-1999; 99US-0129845.
 PR 19-APR-1999; 99US-0130077.
 PR 21-APR-1999; 99US-0130449.
 PR 23-APR-1999; 99US-0130510.
 PR 28-APR-1999; 99US-0130891.
 PR 30-APR-1999; 99US-0132048.
 PR 04-MAY-1999; 99US-0132407.
 PR 05-MAY-1999; 99US-0132485.
 PR 06-MAY-1999; 99US-0132486.
 PR 06-MAY-1999; 99US-0132487.
 PR 07-MAY-1999; 99US-0132863.
 PR 11-MAY-1999; 99US-0134256.
 PR 14-MAY-1999; 99US-0134218.

PR 14-MAY-1999; 99US-0134219.
 PR 14-MAY-1999; 99US-0134221.
 PR 14-MAY-1999; 99US-0134370.
 PR 18-MAY-1999; 99US-0134768.
 PR 19-MAY-1999; 99US-0134941.
 PR 20-MAY-1999; 99US-0135124.
 PR 21-MAY-1999; 99US-0135353.
 PR 24-MAY-1999; 99US-0136021.
 PR 25-MAY-1999; 99US-0136362.
 PR 27-MAY-1999; 99US-0136782.
 PR 28-MAY-1999; 99US-0137222.
 PR 01-JUN-1999; 99US-0137528.
 PR 03-JUN-1999; 99US-0137724.
 PR 07-JUN-1999; 99US-0138094.
 PR 08-JUN-1999; 99US-0138540.
 PR 10-JUN-1999; 99US-0138847.
 PR 14-JUN-1999; 99US-0139119.
 PR 16-JUN-1999; 99US-0139452.
 PR 17-JUN-1999; 99US-0139453.
 PR 18-JUN-1999; 99US-0139454.
 PR 18-JUN-1999; 99US-0139455.
 PR 18-JUN-1999; 99US-0139456.
 PR 18-JUN-1999; 99US-0139457.
 PR 18-JUN-1999; 99US-0139458.
 PR 18-JUN-1999; 99US-0139459.
 PR 18-JUN-1999; 99US-0139460.
 PR 18-JUN-1999; 99US-0139461.
 PR 18-JUN-1999; 99US-0139463.
 PR 18-JUN-1999; 99US-0139750.
 PR 21-JUN-1999; 99US-0139763.
 PR 22-JUN-1999; 99US-0139817.
 PR 23-JUN-1999; 99US-0139869.
 PR 23-JUN-1999; 99US-0140351.
 PR 24-JUN-1999; 99US-0140354.
 PR 28-JUN-1999; 99US-0140695.
 PR 29-JUN-1999; 99US-0140823.
 PR 30-JUN-1999; 99US-0140991.
 PR 01-JUL-1999; 99US-0141287.
 PR 01-JUL-1999; 99US-0141842.
 PR 02-JUL-1999; 99US-0142154.
 PR 06-JUL-1999; 99US-0142055.
 PR 08-JUL-1999; 99US-0142803.
 PR 09-JUL-1999; 99US-0142920.
 PR 12-JUL-1999; 99US-0142977.
 PR 13-JUL-1999; 99US-0143524.
 PR 14-JUL-1999; 99US-0143624.
 PR 15-JUL-1999; 99US-0144005.
 PR 16-JUL-1999; 99US-0144085.
 PR 19-JUL-1999; 99US-0144086.
 PR 19-JUL-1999; 99US-0144325.
 PR 19-JUL-1999; 99US-0144331.
 PR 19-JUL-1999; 99US-0144332.
 PR 19-JUL-1999; 99US-0144333.
 PR 19-JUL-1999; 99US-0144334.
 PR 19-JUL-1999; 99US-0144335.
 PR 20-JUL-1999; 99US-0144352.
 PR 20-JUL-1999; 99US-0144633.
 PR 21-JUL-1999; 99US-0144634.
 PR 21-JUL-1999; 99US-0144884.
 PR 21-JUL-1999; 99US-0145086.
 PR 22-JUL-1999; 99US-0145088.
 PR 22-JUL-1999; 99US-0145085.
 PR 22-JUL-1999; 99US-0145087.
 PR 23-JUL-1999; 99US-0145089.
 PR 23-JUL-1999; 99US-0145192.
 PR 23-JUL-1999; 99US-0145218.
 PR 23-JUL-1999; 99US-0145224.

PR 26-JUL-1999; 99US-0145276.
 PR 27-JUL-1999; 99US-0145913.
 PR 27-JUL-1999; 99US-0145918.
 PR 27-JUL-1999; 99US-0145919.
 PR 28-JUL-1999; 99US-0145951.
 PR 02-AUG-1999; 99US-0146386.
 PR 02-AUG-1999; 99US-0146389.
 PR 03-AUG-1999; 99US-0146389.
 PR 04-AUG-1999; 99US-0147038.
 PR 04-AUG-1999; 99US-0147204.
 PR 05-AUG-1999; 99US-0147302.
 PR 05-AUG-1999; 99US-0147260.
 PR 06-AUG-1999; 99US-0147303.
 PR 06-AUG-1999; 99US-0147416.
 PR 09-AUG-1999; 99US-0147493.
 PR 10-AUG-1999; 99US-0147938.
 PR 11-AUG-1999; 99US-0148171.
 PR 12-AUG-1999; 99US-0148319.
 PR 13-AUG-1999; 99US-0148341.
 PR 13-AUG-1999; 99US-0148565.
 PR 16-AUG-1999; 99US-0149368.
 PR 17-AUG-1999; 99US-0149175.
 PR 18-AUG-1999; 99US-0149426.
 PR 20-AUG-1999; 99US-0149722.
 PR 20-AUG-1999; 99US-0149723.
 PR 23-AUG-1999; 99US-0149929.
 PR 23-AUG-1999; 99US-0149902.
 PR 23-AUG-1999; 99US-0150566.
 PR 26-AUG-1999; 99US-0150884.
 PR 27-AUG-1999; 99US-0151065.
 PR 27-AUG-1999; 99US-0151066.
 PR 30-AUG-1999; 99US-0151080.
 PR 31-AUG-1999; 99US-0151303.
 PR 01-SEP-1999; 99US-0151330.
 PR 07-SEP-1999; 99US-0151330.
 PR 10-SEP-1999; 99US-0152363.
 PR 13-SEP-1999; 99US-0153070.
 PR 15-SEP-1999; 99US-0153758.
 PR 16-SEP-1999; 99US-0154018.
 PR 20-SEP-1999; 99US-0154039.
 PR 22-SEP-1999; 99US-0154779.
 PR 23-SEP-1999; 99US-0155139.
 PR 28-SEP-1999; 99US-0155486.
 PR 29-SEP-1999; 99US-0155659.
 PR 04-OCT-1999; 99US-0156458.
 PR 05-OCT-1999; 99US-0157117.
 PR 06-OCT-1999; 99US-0157753.
 PR 07-OCT-1999; 99US-0157863.
 PR 08-OCT-1999; 99US-0158029.
 PR 12-OCT-1999; 99US-0158232.
 PR 13-OCT-1999; 99US-0158369.
 PR 13-OCT-1999; 99US-0159293.
 PR 13-OCT-1999; 99US-0159294.
 PR 14-OCT-1999; 99US-0159295.
 PR 14-OCT-1999; 99US-0159329.
 PR 14-OCT-1999; 99US-0159330.
 PR 14-OCT-1999; 99US-0159331.
 PR 14-OCT-1999; 99US-0159637.
 PR 18-OCT-1999; 99US-0159584.
 PR 21-OCT-1999; 99US-0160741.
 PR 21-OCT-1999; 99US-0160767.
 PR 21-OCT-1999; 99US-0160770.
 PR 21-OCT-1999; 99US-0160776.
 PR 21-OCT-1999; 99US-0160814.
 PR 22-OCT-1999; 99US-0160815.
 PR 22-OCT-1999; 99US-0160980.
 PR 22-OCT-1999; 99US-0160981.
 PR 22-OCT-1999; 99US-0160989.

PR	25-OCT-1999;	99US-0161404.
PR	25-OCT-1999;	99US-0161405.
PR	25-OCT-1999;	99US-0161406.
PR	25-OCT-1999;	99US-0161359.
PR	25-OCT-1999;	99US-0161360.
PR	26-OCT-1999;	99US-0161361.
PR	26-OCT-1999;	99US-0161920.
PR	28-OCT-1999;	99US-0161993.
PR	28-OCT-1999;	99US-0161993.
PR	29-OCT-1999;	99US-0162142.

Query Match	57.58;	Score 42;	0;
Similarity	53.88;	Pred. No. 40;	0;
Best Local	7;	Indels	4;
Matches	7;	Mismatches	2;
		Conservative	0;

```
QY      1 0ASIEGRTAPQP 13
      | : ||: |||
Db      390 QEGLAGRTPEQP 402
```

RESULT 8
AAG3806
000006 standard: Protein; 493 AA.

ID AAG3806
 XX AAG3806;
 AC
 XX
 DT 17-OCT-2000 (first entry)
 fragment SEQ ID NO: 27249.

XX Arabidopsis thaliana protein region
DE
XX
XX Protein identification; signal transduction pathway; metabolic pathway; promoter;
KM hybridisation assay; genetic mapping; gene expression control;
KM termination sequence.

XX Arabidopsis thaliana.
OS
XX
PN EPI033405-A2.

XX 06-SEP-2000.
PD
XX
XX 25-FEB-2000; 2000EP-0301439.
PF
XX 25-FEB-1999; 9905-0121825.
XX

PR 25-MAR-1999; 9905-0126785.
PR 23-MAR-1999; 9905-0126264.
PR 09-MAR-1999; 9905-0125788.
PR 05-MAR-1999; 9905-0123548.
PR 25-MAR-1999; 9905-0123100.

PR 29-MAR-1999; 9905-0127462.
PR 01-APR-1999; 9905-0128234.
PR 06-APR-1999; 9905-0128714.
PR 08-APR-1999; 9905-0129845.
PR 16-APR-1999; 9905-0130077.

PR 19-APR-1999; 99US-0130449.
PR 21-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130691.
PR 23-APR-1999; 99US-0131449.
28-APR-1999; 99US-0131449.

PR 26-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 30-APR-1999; 99US-0132484.
PR 04-MAY-1999; 99US-0132485.
PR 05-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.

PR 06-MAY-1999; 9905-0132467.
PR 07-MAY-1999; 9905-0132863.
PR 11-MAY-1999; 9905-0134256.
PR 14-MAY-1999; 9905-0134218.
PR 14-MAY-1999; 9905-0134219.
PR 14-MAY-1999; 9905-0134219.

PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 14-MAY-1999; 99US-0134768.
PR 18-MAY-1999; 99US-0134941.
PR 1900; 99US-0134941.

PR 19-MAY-1995; 99US-0135124.
PR 20-MAY-1999; 99US-0135353.
PR 21-MAY-1999;

PR 24-MAY-1999; 99US-0135629.
25-MAY-1999; 99US-0136021.
26-MAY-1999; 99US-0136022.

PR 27-MAY-1999; 99US-0136352.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 01-JUN-1999; 99US-0137528.
PR 03-JUN-1999; 99US-0137502.

PR 04-JUN-1999; 99US-013772.
PR 07-JUN-1999; 99US-013809.
PR 08-JUN-1999; 99US-013854.
PR 10-JUN-1999; 99US-013887.

PR 10-JUN-1999; 99US-0139119.
PR 14-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 16-JUN-1999; 99US-0139492.
PR 1999; 99US-0139492.

PR 17-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.

PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.

PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 18-JUN-1999; 99US-0139817.

PR 21-JUN-1999; 99US-0139899.
PR 22-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 23-JUN-1999; 99US-0140695.

PR	24-JUN-1999;	99US-0140823.
PR	28-JUN-1999;	99US-0140991.
PR	29-JUN-1999;	99US-0141287.
PR	30-JUN-1999;	99US-0141842.
PR	01-JUL-1999;	99US-0141842.

PR 01-JUL-1999; 99US-0142154.
PR 01-JUL-1999; 99US-0142055.
PR 02-JUL-1999; 99US-0142390.
PR 06-JUL-1999; 99US-0142803.
PR 08-JUL-1999; 99US-0142820.

PR	00-JUL-1999;	99US-0142520.
PR	09-JUL-1999;	99US-0142977.
PR	12-JUL-1999;	99US-0143542.
PR	13-JUL-1999;	99US-0143624.
PR	14-JUL-1999;	99US-0144005.

PR 15-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 16-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 1999; 99US-0144332.

PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 19-JUL-1999; 99US-0144335.

PR 19-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 20-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.

PR 21-JUL-1999; 99US-0145088.
PR 21-JUL-1999; 99US-0145088.
PR 21-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.

PR 22-JUL-1999; 99US-0145192.
PR 22-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 23-JUL-1999; 99US-0145224.

PR	23-JUL-1999;	99US-0145276.
PR	26-JUL-1999;	99US-0145913.
PR	27-JUL-1999;	99US-0145918.
PR	27-JUL-1999;	99US-0145919.
PR	27-JUL-1999;	99US-0145919.

PR 27-JUL-1999; 99US-0145951.
PR 28-JUL-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999;

PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 06-AUG-1999; 99US-0147192.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 09-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 10-AUG-1999; 99US-0147935.
PR 11-AUG-1999; 99US-0148171.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 23-AUG-1999; 99US-0149923.
PR 23-AUG-1999; 99US-0149902.
PR 25-AUG-1999; 99US-0149930.
PR 26-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 30-AUG-1999; 99US-0151080.
PR 31-AUG-1999; 99US-0151303.
PR 01-SEP-1999; 99US-0151438.
PR 07-SEP-1999; 99US-0151930.
PR 10-SEP-1999; 99US-0152363.
PR 13-SEP-1999; 99US-0153070.
PR 15-SEP-1999; 99US-0153755.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154039.
PR 22-SEP-1999; 99US-0154779.
PR 23-SEP-1999; 99US-0155139.
PR 24-SEP-1999; 99US-0155486.
PR 28-SEP-1999; 99US-0155659.
PR 29-SEP-1999; 99US-0156458.
PR 04-OCT-1999; 99US-0156596.
PR 05-OCT-1999; 99US-0157117.
PR 06-OCT-1999; 99US-0157753.
PR 07-OCT-1999; 99US-0157865.
PR 08-OCT-1999; 99US-0158023.
PR 12-OCT-1999; 99US-0158232.
PR 13-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 14-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 18-OCT-1999; 99US-0159638.
PR 21-OCT-1999; 99US-0159844.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 26-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 28-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.

PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match
Best Local Similarity 57.5%; Score 42; DB 21; Length 493;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 1 QASIBRYTAPOP 13
Db 409 QESLAGREFTEPOP 421

RESULT 9
AA013603
ID AA013603 standard; Protein; 69 AA.
XX
XX AA013603;
AC
DT 06-NOV-2001 (first entry)
DE
XX
XX Human polypeptide SEQ ID NO 27495.
KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
XX nervous system disorders; arthritis; inflammation;
OS Homo sapiens.

PN WO200164835-A2.
XX
XX 07-SEP-2001.
PD
XX
XX 26-FEB-2001; 2001WO-US04927.
PE
XX
XX 28-FEB-2000; 2000US-0515126.
PR 18-MAY-2000; 2000US-0577409.
XX
XX (HYSE-) HYSEQ INC.
PA
XX
XX Tang YT, Liu C, Dymnac RT;
PI WPI; 2001-514838/56.
XX
XX N-PSDB; AA193534.

Isolated nucleic acids and polypeptides, useful for preventing
diagnosing and treating e.g. leukaemia, inflammation and immune
disorders.
PS Claim 20; SEQ ID NO 27495; 1399pp + Sequence Listing; English.
XX
XX The invention relates to human polynucleotides (AA179941-AA193841) and
XX the encoded proteins (AA000010-AA013910) that exhibit activity elating to
XX cytokine, cell proliferation or cell differentiation or which may induce
XX production of other cytokines in other cell populations. The
XX polynucleotides and polypeptides are useful in gene therapy, vaccines or
XX peptide therapy. The polypeptides have various cytokine-like activities,
XX e.g. stem cell growth factor activity, haematopoiesis regulating
XX activity, tissue growth factor activity, immunomodulatory activity and
XX treatment of cancer, leukaemia, nervous system disorders, arthritis and
XX inflammation.
XX Note: The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
SQ Sequence 69 AA;

Query Match
Best Local Similarity 53.3%; Score 41.5; DB 22; Length 69;
Matches 8; Conservative 4; Mismatches 2; Indels 1; Gaps 1;

PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;
 DR WPI; 2001-48897/53.
 XX

PT Human genome-derived single exon nucleic acid probes useful for
 PT analyzing gene expression in human placenta -
 XX
 PS Claim 27; SEQ ID No 33581; 654pp; English.

CC The present invention relates to single exon nucleic acid probes (SEMP;
 CC see A131315-A157346). The present sequence is a peptide encoded by one
 CC such probe. The probes are useful for producing a microarray for
 CC predicting, measuring and displaying gene expression in samples derived
 CC from human placenta. The probes are useful for antenatal diagnosis of
 CC human genetic disorders.
 CC
 SQ Sequence 49 AA;

Query Match 56.2%; Score 41; DB 22; Length 49;
 Best Local Similarity 60.0%; Pred. No. 4.7;
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 4 IEGRYTAPOP 13
 :||:|:|:|
 DB 20 VEGRTVPRP 29

RESULT 13

ID ABG42939 standard; Peptide; 49 AA.

AC ABG42939;

DT 19-AUG-2002 (first entry)

DE Human peptide encoded by genome-derived single exon probe SEQ ID 32604.

XX Human; single exon probe; asthma; lung cancer; COPD; ILD;
 KW chronic obstructive pulmonary disease; interstitial lung disease;
 KW familial idiopathic pulmonary fibrosis; neurofibromatosis;
 KW tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
 KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
 KW pulmonary histiocytosis; lymphangioleiomyomatosis; Karagener syndrome;
 KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
 KW primary ciliary dyskinesia; pulmonary hypertension;
 KW hyaline membrane disease.
 XX
 OS Homo sapiens.

PN WO200186003-A2.
 PD 15-NOV-2001.
 XX

PF 30-JAN-2001; 2001WO-US00665.
 XX

PR 04-FEB-2000; 2000US-180312P.
 PR 26-MAY-2000; 2000US-207456P.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-234687P.
 PR 27-SEP-2000; 2000US-236359P.
 PR 04-OCT-2000; 2000GB-0024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;
 DR WPI; 2002-114183/15.
 XX

PT Spatially-addressable set of single exon nucleic acid probes, used to
 PT measure gene expression in human lung samples -
 XX
 PS Claim 27; SEQ ID No 32604; 634pp; English.

CC The invention relates to a spatially-addressable set of single exon
 CC nucleic acid probes for measuring gene expression in a sample derived
 CC from human lung comprising single exon nucleic acid probes having one of
 CC 12614 nucleic acid sequences mentioned in the specification, or their
 CC complements or the 12387 open reading frames derived from the 12614
 CC probes. Also included are a microarray comprising the novel set of
 CC probes; the novel set of probes which hybridize at high stringency to a
 CC sample derived from human lung; measuring gene expression in a
 CC collection of detectably labeled nucleic acids derived from human lung
 CC mRNA; and (b) measuring the label detectably bound to each probe of
 CC the array; identifying exons in a eukaryotic genome, comprising
 CC (a) algorithmically predicting at least one exon from genomic sequences
 CC labeled nucleic acids from eukaryotic lung mRNA, to a single exon probe,
 CC having a fragment identical to the predicted exon, the probe is included
 CC in the above mentioned microarray; assigning exons to a single gene,
 CC comprising (a) identifying exons from genomic sequence by the method
 CC above and (b) measuring the expression of each of the exons in several
 CC tissues and/or cell types using hybridization to a single exon
 CC microarrays having a probe with the exon, where a common pattern of
 CC expression of the exons in the tissues and/or cell types indicates that
 CC the exons should be assigned to a single gene; a peptide comprising one
 CC of 12011 sequences, mentioned in the specification, or encoded by the
 CC probes/open reading frames (ORF). The probes are used for gene
 CC expression analysis, and for identifying exons in a gene, particularly
 CC such as asthma, lung cancer, chronic obstructive pulmonary disease
 CC (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary
 CC fibrosis, neurofibromatosis, tuberous sclerosis, Gaucher's disease,
 CC Niemann-Pick disease, Hermansky-Pudlak syndrome, sarcoidosis, pulmonary
 CC haemosiderosis, pulmonary histiocytosis, lymphangioleiomyomatosis,
 CC pulmonary alveolar proteinosis, Karagener syndrome, fibrocystic
 CC pulmonary dysplasia, primary ciliary dyskinesia, pulmonary hypertension
 CC and hyaline membrane disease. The present sequence is a peptide/protein
 CC encoded by a single exon probe of the invention.
 CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic
 CC format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.

SQ Sequence 49 AA;

Query Match 56.2%; Score 41; DB 23; Length 49;
 Best Local Similarity 60.0%; Pred. No. 4.7;
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 4 IEGRYTAPOP 13
 :||:|:|:|
 DB 20 VEGRTVPRP 29

RESULT 14
 ABG07274
 ID ABG07274 standard; Protein; 186 AA.

AC ABG07274;

DT 13-FEB-2002 (first entry)

DE Novel human diagnostic protein #7265.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.

XX OS Homo sapiens.
 XX PN WO200175067-A2.
 XX PD 11-OCT-2001.
 XX PF 30-MAR-2001; 2001WO-US08631.
 XX PR 31-MAR-2000; 2000US-0540217.
 XX PR 23-AUG-2000; 2000US-0649167.
 XX PA (HYSE-) HYSEQ INC.
 XX PI Drmanac RT, Liu C, Tang YT;
 XX DR WPI; 2001-639362/73.
 XX DR N-PSDB; AAS71461.
 XX PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity.
 XX PS Claim 20; SEQ ID No 37633; 103pp; English.
 XX CC The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX SQ Sequence 186 AA;
 OY 3 SIEGRYAPQ 14
 DB 166 NVNGRDFAPQ 177
 Query Match 56.2%; Score 41; DB 22; Length 186;
 Best Local Similarity 58.3%; Pred. No. 21;
 Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
 RESULT 15
 AAY22919
 ID AAY22919 standard; peptide; 11 AA.
 XX AC AAY22919;
 XX DT 19-AUG-1999 (first entry)
 XX DE Galactose oxidase linker used to make proteinase inhibitor fusions.
 XX KW Peptide linker; fusion protein; pathogen resistance; pathogen tolerance;
 KW plant; transgene; proteinase inhibitor; Arabidopsis thaliana;
 KW parasitic nematode.
 XX

OS Synthetic.
 XX PN WO9928484-A1.
 XX PD 10-JUN-1999.
 XX PF 01-DEC-1998; 98WO-EP07792.
 XX PR 03-DEC-1997; 97GB-0025556.
 XX PA (NOVS) NOVARTIS AG.
 XX PI Atkinson HJ, McPherson MJ, Urwin PE;
 XX DR WPI; 1999-385387/32.
 XX PT Proteinase inhibitor fusion proteins
 XX PS Claim 9; Page 23; 39pp; English.
 XX CC The present sequence represents a peptide linker used in the fusion
 CC proteins of the invention. The specification describes a method for
 CC improving pathogen resistance or tolerance of a plant. The method
 CC comprises transformation of the plant with a transgene encoding a
 CC fusion protein which comprises two or more protein or domains that
 CC are capable of improving pathogen resistance or tolerance when
 CC expressed on their own. Specifically, two distinct proteinase
 CC inhibitors are co-delivered, as a fusion, to Arabidopsis thaliana. The
 CC proteins or domains are connected by a peptide linker. The method is
 CC used to improve pathogen resistance or tolerance of a plant and its
 CC descendants, especially against parasitic nematode attack.
 XX SQ Sequence 11 AA;
 OY 1 QASIEGRYAPQ 14
 DB 1 QAS---STYAPQ 11
 Query Match 55.5%; Score 40.5; DB 20; Length 11;
 Best Local Similarity 71.4%; Pred. No. 1.1;
 Matches 10; Conservative 0; Mismatches 1; Indels 3; Gaps 1;
 Search completed: January 13, 2003, 09:55:42
 Job time : 32.9487 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 13, 2003, 09:53:20 ; Search time 53.1282 Seconds

(without alignments)
54.296 Million cell updates/sec

Title: US-09-554-941-1
Perfect score: 70
Sequence: 1 VILGVPKIQIEG 14

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: SPREMBL_21:*
2: sp_archaea:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phase:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp Vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	65	92.9	75	10	Q9SP23 medicago sa
2	53	75.7	74	10	Q23812 hordeum vul
3	46	65.7	75	10	Q9LUX2 pyrus pyrif
4	43	61.4	77	10	Q93X22 quercus sub
5	43	61.4	79	10	Q9AUI6 typha latif
6	43	61.4	79	10	Q944W1 typha latif
7	42	60.0	81	10	Q94187 atropa bell
8	42	60.0	602	10	Q9SC90 medicago tr
9	41	58.6	69	10	Q82562 elaeagnus u
10	41	58.6	80	10	Q9FR40 petunia hyb
11	41	58.6	453	4	Q8WV65 homo sapien
12	40	57.1	54	16	Q06352 mycobacteri
13	40	57.1	79	10	Q80335 citrus unsh
14	40	57.1	82	10	Q94198 atropa bell
15	40	57.1	201	16	Q92R37 rhizobium m
16	40	57.1	212	2	Q9EYB2 escherichia

17	40	57.1	215	16	Q8XSE1 escherichia
18	40	57.1	456	16	Q8ZGF4 versinia pe
19	39	55.7	76	10	Q9M4N0 persia amer
20	39	55.7	360	17	Q9HMA0 halobacteri
21	39	55.7	441	13	Q12969 gallus gall
22	39	55.7	495	10	Q9STM3 solanum tub
23	39	55.7	499	10	Q941V1 arabidopsis
24	39	55.7	508	10	Q80874 arabidopsis
25	39	55.7	512	10	Q9LML0 bacillus ha
26	39	55.7	813	16	Q9KGS2 ipomoea bat
27	38	54.3	66	10	Q9ZSC6 methanosarc
28	38	54.3	115	17	Q8TKX4 rhizobium l
29	38	54.3	167	16	Q98K40 archaeglob
30	38	54.3	237	17	Q59554 pyrococcus
31	38	54.3	334	4	Q29818 archaeoglob
32	38	54.3	343	17	Q8TE01 actinobact
33	38	54.3	368	2	Q9LAD4 clostridium
34	38	54.3	396	16	Q8XN53 bradyrhizob
35	38	54.3	430	16	Q97HAI infectious
36	38	54.3	556	2	Q45232 P90736
37	38	54.3	616	12	Q91BR3 P90736
38	38	54.3	616	12	Q91BR3 P90736
39	38	54.3	616	12	Q91BR3 P90736
40	38	54.3	2632	5	P90736 P90736
41	37	52.9	260	16	Q987CO P90736
42	37	52.9	350	16	Q8YCX3 P90736
43	37	52.9	371	10	Q9SUY4 P90736
44	37	52.9	382	2	Q9ZB53 P90736
45	37	52.9	392	2	Q9ZB53 P90736

ALIGNMENTS

RESULT 1
Q9SP23 PRELIMINARY: PRT: 75 AA.
AC Q9SP23; 01-MAY-2000 (TREMBL) 13, Created)
DT 01-MAY-2000 (TREMBL) 13, Last sequence update)
DT 01-DEC-2001 (TREMBL) 19, Last annotation update)
DE Type 1 metallothionein.
GN MET1.
OS Medicago sativa (Alfalfa).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabiales; Fabaceae; Papilionoideae; Trifoliales; Medicago.
OX NCBI_TaxID=3879;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. LAHONTAN; TISSUE=ROOT;
RX PubMed=11448760;
RA Potenza C., Thomas S.H., Sengupta-Gopalan C.;
RT "Genes induced during early response to Meloidogyne incognita in roots
of resistant and susceptible alfalfa cultivars.";
RL Plant Sci. 161:289-299(2001).
DR EMBL: AF189766; AAF04584.1;
DR InterPro: IPR000347; Metallothion_15.
DR Pfam: PF01439; Metallothion_15; 1.
DR ProDom: PD001611; Metallothion_15; 1.
SQ SEQUENCE 75 AA; 7537 MW; 7872EF0F02BD22FB CRC64;
Query Match Score 65; DB 10; Length 75;
Best Local Similarity 92.9%; Pred. No. 0.00038;
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 VILGVPKIQIEG 14
|||||
Db 36 VILGVPKIQIEG 49

RESULT 2
Q23812

ID 023812 PRELIMINARY; PRT; 74 AA.
 AC 023812;
 DT 01-JAN-1998 (TREMBlrel. 05, Created)
 DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE Plant metallothionein-like protein.
 GN ID1.
 OS Hordeum vulgare (Barley).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
 OC Triticeae; Hordeum.
 OC NCBI_TaxID=4513;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=VAR. NK1558;
 RA Nakanishi H., Okumura N., Kanegae R., Umehara Y., Nishizawa N.,
 Mori S.;
 RT "A plant metallothionein-like gene from iron deficiency barley
 roots.";
 RL Submitted (MAY-1995) to the EMBL/GenBank/DBJ databases.
 DR EMBL; D50641; BA23628.1;
 DR InterPro: IPR000347; Metallothion_15.
 DR Pfam: PF01439; Metallothio_2; 1.
 DR ProDom: PD001611; Metallothion_15; 1.
 SQ SEQUENCE 74 AA; 7409 MW; F86C8CC3ED58B1C6 CRC64;

Query Match 75.7%; Score 53; DB 10; Length 74;
 Best Local Similarity 69.2%; Pred. No. 0.054;
 Matches 9; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 VILGVPKIOFEG 13
 :|||||
 Db 37 IIVGVAPKMFEG 49

RESULT 3
 ID 09LUX2 PRELIMINARY; PRT; 75 AA.
 AC 09LUX2;
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
 DE Metallothionein-like protein.
 GN PPR016.
 OS Pyrus pyrifolia (Japanese pear) (Pyrus serotina).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC Eurosid 1; Rosales; Rosaceae; Maloideae; Pyrus.
 OC NCBI_TaxID=3767;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. KIKUSUI; TISSUE=FRUIT;
 RA Itai A., Tanaka T., Tanabe K., Tamura F.;
 RT "Pyrus pyrifolia ripening associated mRNA, clone PPR016.";
 RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB021790; BA96449.1;
 DR InterPro: IPR000347; Metallothion_15.
 DR Pfam: PF01439; Metallothio_2; 1.
 DR ProDom: PD001611; Metallothion_15; 1.
 SQ SEQUENCE 75 AA; 7831 MW; A772B9B99367755A CRC64;

Query Match 65.7%; Score 46; DB 10; Length 75;
 Best Local Similarity 57.1%; Pred. No. 0.09;
 Matches 8; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

OY 1 VILGVPKIOFEG 14
 :|||||
 Db 36 IIVGVAPKMFEG 49

RESULT 4
 ID 093X22 PRELIMINARY; PRT; 77 AA.

AC 093X22;
 DT 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
 DE Metallothionein-like protein.
 GN MT.
 OS Quercus suber (Cork oak).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC Eurosid 1; Fagales; Fagaceae; Quercus.
 OC NCBI_TaxID=58331;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=PERIDERM;
 RA Mir G.;
 RT "Cork metallothionein.";
 RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AJ277599; CAC39481.2;
 DR InterPro: IPR000347; Metallothion_15.
 DR Pfam: PF01439; Metallothio_2; 1.
 DR ProDom: PD001611; Metallothion_15; 1.
 SQ SEQUENCE 77 AA; 7673 MW; C867A295A9DAB0A5 CRC64;

Query Match 61.4%; Score 43; DB 10; Length 77;
 Best Local Similarity 57.1%; Pred. No. 3.5;
 Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 1 VILGVPKIOFEG 14
 :|||||
 Db 39 IIVGVAPKMFEG 52

RESULT 5
 ID 09AUI6 PRELIMINARY; PRT; 79 AA.
 AC 09AUI6;
 DT 01-JUN-2001 (TREMBlrel. 17, Created)
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
 DT 01-OCT-2001 (TREMBlrel. 18, Last annotation update)
 DE Metallothionein-like protein.
 GN Typha latifolia (Common cattail).
 OS Typha latifolia (Common cattail).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Typhaceae; Typha.
 OC NCBI_TaxID=4733;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Zhang Y.W., Ma M., Tam F.Y., Wong Y.S.;
 RT "Cloning and sequencing of metallothionein (MT) cDNA from a heavy
 metal tolerant plant, Typha latifolia.";
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF279655; AA28022.1;
 DR InterPro: IPR000347; Metallothion_15.
 DR Pfam: PF01439; Metallothio_2; 1.
 DR ProDom: PD001611; Metallothion_15; 1.
 SQ SEQUENCE 79 AA; 7954 MW; B81C153D3E4EC2D CRC64;

Query Match 61.4%; Score 43; DB 10; Length 79;
 Best Local Similarity 64.3%; Pred. No. 3.6;
 Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 1 VILGVPKIOFEG 14
 :|||||
 Db 40 IIVGVAPKMFEG 53

RESULT 6
 ID 094W1 PRELIMINARY; PRT; 79 AA.
 AC 094W1;
 DT 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
 DE Type 2 metallothionein-like protein.

Mon Jan 13 10:35:10 2003

us-09-554-941-1.rsp

Page 3

OS Typha latifolia (Common cattail).
OC Spermatophyta: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta: Magnoliophyta; Liliopsida; Poales; Typhaceae; Typha.
OX NCBI_TaxID=4733;
[1]
RN SEQUENCE FROM N.A.
RP Zhang Y.-W., Tam F.-Y., Wong Y.-S.;
RA "Cloning and characterization of type 2 metallothionein-like gene from
RT Typha latifolia."
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF419840; AL09705.1; Metallothion_15.
DR InterPro: IPR000347; Metallothion_15; 1.
DR Pfam: PF01439; Metallothion_15; 1.
DR ProDom: PD001611; Metallothion_15; 1.
DR SEQUENCE 79 AA; 8053 MW; B8090030C6F1EC2D CRC64;
SQ
Query Match 61.4%; Score 43; DB 10; Length 79;
Best Local Similarity 64.3%; Pred. No. 3.6; 4; Indels 0; Gaps 0;
Matches 9; Conservative 1; Mismatches 0;
QY 1 VILGVGPAKIQFEG 14
DB 40 MILGVAPKGNFEG 53
RESULF 7
Q94187 PRELIMINARY; PRT; 81 AA.
ID Q94187
AC 094187
DT 01-DEC-2001 (TREMBLER, 19, Created)
DT 01-DEC-2001 (TREMBLER, 19, Last sequence update)
DT 01-JUN-2002 (TREMBLER, 21, Last annotation update)
DE Putative metallothionein-like protein type 2B.
DE Atropa belladonna (belladonna) (Deadly nightshade). Tracheophyta;
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Atropa.
OX NCBI_TaxID=33113;
[1]
RN SEQUENCE FROM N.A.
RP Nouar E.;
RA "Differential gene expression in Atropa belladonna leafy gall."
RT Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
RL [2]
RN SEQUENCE FROM N.A.
RP Nouar E.;
RA Thesis (2001), Department of Plant Biotechnology,
RL Universite Libre de Bruxelles, Brussels, Belgium.
RL EMBL: AJ309387; CAC40757.1;
DR InterPro: IPR000347; Metallothion_15.
DR Pfam: PF01439; Metallothion_2; 1.
DR ProDom: PD001611; Metallothion_15; 1.
DR SEQUENCE 81 AA; 8294 MW; 1CE23612CDBB9F4 CRC64;
SQ
Query Match 60.0%; Score 42; DB 10; Length 81;
Best Local Similarity 53.8%; Pred. No. 5.6; 3; Indels 0; Gaps 0;
Matches 7; Conservative 3; Mismatches 0;
QY 1 VILGVGPAKIQFEG 13
DB 40 LVLVGVPKSTFSD 52
RESULT 8
ID Q9SC90 PRELIMINARY; PRT; 602 AA.
AC 09SC90
DT 01-MAY-2000 (TREMBLER, 13, Created)
DT 01-OCT-2000 (TREMBLER, 15, Last sequence update)
DT 01-MAR-2002 (TREMBLER, 20, Last annotation update)
DE Pectin methyl-esterase PER precursor (PC 3.1.1.11).
GN PER.
OS Medicago truncatula (Barrel medic).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Trifoliaceae; Medicago.
OX NCBI_TaxID=3880;
[1]
RN SEQUENCE FROM N.A.
RP STRAIN-CV, JEMALONG;
RA Perez-Hormaeche J., Kondorosi A., Palomares A.J., Ratet P.;
RT "MPER, a new pectin methyl-esterase gene from Medicago truncatula
involved in the infection process of root nodules by the symbiotic
host Sinorhizobium meliloti."
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
[2]
RN SEQUENCE FROM N.A.
RP STRAIN-CV, JEMALONG;
RC Perez-Hormaeche J., Rodriguez-Llorente I.D., Kondorosi A.,
RA Palomares A.J., Ratet P.;
RT "A genomic cluster of pectin methyl-esterase genes strongly suggest a
common evolutionary origin for isoforms induced in pollen and root
nodules development."
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ249611; CAB65290.2;
DR InterPro: IPR000070; Pectinesterase.
DR InterPro: IPR003880; Pantane attach.
DR Pfam: PF01095; Pectinesterase; 1.
DR PROSITE: PS00800; PECTINESTERASE; 1.
DR PROSITE: PS00012; PHOSPHOPANTETHEINE; UNKNOWN_1.
DR Hydrolase; signal.
KW SIGNAL.
FT SEQUENCE 602 AA; 65003 MW; FB827E825FE62B2F CRC64;
SQ
Query Match 60.0%; Score 42; DB 10; Length 602;
Best Local Similarity 61.5%; Pred. No. 45; 3; Indels 0; Gaps 0;
Matches 8; Conservative 2; Mismatches 0;
QY 2 ILGVGPAKIQFEG 14
DB 284 ILGVGPAKIQFEG 296
RESULT 9
ID 082562 PRELIMINARY; PRT; 69 AA.
AC 082562
DT 01-NOV-1998 (TREMBLER, 08, Created)
DT 01-NOV-1998 (TREMBLER, 08, Last sequence update)
DT 01-MAR-2002 (TREMBLER, 20, Last annotation update)
DE Metallothionein homolog.
DE Eleagnus umbellata (Autumn olive).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Rosales; Elaeagnaceae; Elaeagnus.
OX NCBI_TaxID=43233;
[1]
RN SEQUENCE FROM N.A.
RP TISSUE-ROOT NODULE;
RA Kim H.B., An C.S.;
RT "Isolation and characterization of cDNA clone encoding metallothionein
from the root nodule of Eleagnus umbellata."
RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF091514; AAC62105.1;
DR InterPro: IPR000347; Metallothion_15.
DR Pfam: PF01439; Metallothion_2; 1.
DR ProDom: PD001611; Metallothion_15; 1.
DR SEQUENCE 69 AA; 7141 MW; 2B98518FEB08938A CRC64;
SQ
Query Match 58.6%; Score 41; DB 10; Length 69;
Best Local Similarity 61.5%; Pred. No. 7.2; 4; Indels 0; Gaps 0;
Matches 8; Conservative 1; Mismatches 0;
QY 2 ILGVGPAKIQFEG 14
DB 31 VLGVAPKGNFEG 43

RESULT 10

Q9FR40
ID 09FR40 PRELIMINARY; PRT; 80 AA.
AC 09FR40
DT 01-MAR-2001 (TRENBLREL. 16, Created)
DT 01-MAR-2001 (TRENBLREL. 16, Last sequence update)
DE 01-DEC-2001 (TRENBLREL. 19, Last annotation update)
CN Putative metallothionein-like protein.
OS Petunia hybrida (Petunia).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OX NCBI_TaxID=4102;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PETUNIA MULTIFLORA HYBRID, AND CV. PRIMETIME ROSE;
RA Lai C.P., Shaw J.F.;
RT Cloning and characterization of cDNAs for metallothionein-like
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF201384; AAC39645.1;
DR InterPro: IPR000347; Metallothion_15.
DR Pfam: PF01439; Metallothio_2; 1.
DR Prodom: PD001611; Metallothion_15; 1.
SQ SEQUENCE 80 AA; 8084 MW; 4248ADE0C05FACC7 CRC64;

Query Match

Best Local Similarity 58.6%; Score 41; DB 10; Length 80;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 ILGVSPAKIOF 12
Db 40 ILIGVPEKTF 51
:|||||:
:|||||:

RESULT 11

08WV65
ID 08WV65 PRELIMINARY; PRT; 453 AA.
AC 08WV65
DT 01-MAR-2002 (TRENBLREL. 20, Created)
DT 01-MAR-2002 (TRENBLREL. 20, Last sequence update)
OS Hypothetical 50.0 kDa protein.
DE Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN;
RA Strausberg R.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC018705; AAH18705.1;
DR InterPro: IPR001392; Clathrin_med.
DR Pfam: PF00928; Adap_comp_sub; 1.
KW Hypothetical protein.
SQ SEQUENCE 453 AA; 49977 MW; AEDCA8C5AED08B7 CRC64;

Query Match

Best Local Similarity 58.6%; Score 41; DB 4; Length 453;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 3 LGVSPAKIOF 13
Db 400 LGGSPASISFE 410
:|||||:
:|||||:

RESULT 12

006352
ID 006352 PRELIMINARY; PRT; 54 AA.

RESULT 13

080335
ID 080335 PRELIMINARY; PRT; 79 AA.
AC 080335
DT 01-NOV-1998 (TRENBLREL. 08, Created)
DT 01-NOV-1998 (TRENBLREL. 08, Last sequence update)
DE Metallothionein-like protein.
CN R3489 OR MTCY13E12.42 OR MT3593.
OS Mycobacterium tuberculosis.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H37RV;
RX MEDLINE=98295987; PubMed=9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eigmeier K., Gas S., Barry C.E. III, Tekala E.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Pettell T., Gentles S., Hamlin N., Holroyd S.,
RA Hornsby T., Jagels K., Krogh A., Mclean A., Rajandream M.A., Rogers J.,
RA Oliver S., Osborne J., Quail M.A., Rastandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skellton S., Squares S., Squares R.,
RA Sulston J.E., Taylor K., Whitehead S., Barrett B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RL complete genome sequence."
RN Nature 393:537-544(1998).
[2]
RP SEQUENCE FROM N.A.
RC STRAIN=CDC 1551 / OSHKOSH;
RA Felschmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J., Deboy R., Dodson R., Gwinn M.L., Hart D., Hickey E.,
RA Kolonay J.F., Nelson W.C., Umayan L.A., Emdolaeva M.D., Salzberg S.L.,
RA Delcher A., Uitterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA Bishai W.;
RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
RL laboratory strains."
DR EMBL; Z95390; CAB08712.1;
DR EMBL; AE007162; AAK47952.1;
DR TIGR; MT3593;
KW TubercuList; RV3489;
SQ SEQUENCE 54 AA; 5695 MW; 3406A9C3DF1FBC96 CRC64;

Query Match

Best Local Similarity 57.1%; Score 40; DB 16; Length 54;
Matches 6; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

OY 2 ILGVSPAKIO 11
Db 44 MLGSPAKIE 53
:|||||:
:|||||:

RESULT 13

080335
ID 080335 PRELIMINARY; PRT; 79 AA.
AC 080335
DT 01-NOV-1998 (TRENBLREL. 08, Created)
DT 01-NOV-1998 (TRENBLREL. 08, Last sequence update)
DE Metallothionein-like protein.
CN R3489 OR MTCY13E12.42 OR MT3593.
OS Mycobacterium tuberculosis.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=FRUIT;
RX MEDLINE=98267202; PubMed=9602134;
RA Moriyuchi T., Kita M., Hisada S., Endo-Inagaki T., Omura M.;
RT "Characterization of gene repertoires at mature stage of citrus fruits
through random sequencing and analysis of redundant metallothionein-
like genes expressed during fruit development."

RL Gene 211:221-227(1998).
 DR EMBL: AB008100; BAA31561.1; -;
 DR InterPro: IPR000347; Metallothion_15.
 DR Pfam: PF01439; Metallothio_2; 1.
 DR ProDom: PD001611; Metallothion_15; 1.
 DR SEQUENCE 79 AA; 7741 MW; CDDAE5CC41BD36AE CRC64;

Query Match 57.1%; Score 40; DB 10; Length 79;
 Best Local Similarity 50.0%; Pred. No. 13;
 Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 LVIGVPAKIQFEG 14
 :||| | :|||
 Db 42 LVIGVAPVKMHSEG 55

RESULT 14

Q94198 PRELIMINARY; PRT; 82 AA.

AC Q94198; 01-DEC-2001 (TREMblrel. 19, Created)

DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)

DE 01-JUN-2002 (TREMblrel. 21, Last annotation update)

DE Methallothioneine-like protein.

GN MT2B.

OS Atropa belladonna (Belladonna) (Deadly nightshade);

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

OC Asteridae; euasterids I; Solanales; Solanaceae; Atropa.

NCBI_TaxID=33113;

RM [1]

RP SEQUENCE FROM N.A.

RA Nouar E., Marie B., Jaziri M.;

RT "Differential gene expression in *Atropa belladonna* leafy gall.";

RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL: AJ297968; CAC40742.1; -;

DR InterPro: IPR000347; Metallothion_15.

DR Pfam: PF01439; Metallothio_2; 1.

DR ProDom: PD001611; Metallothion_15; 1.

DR SEQUENCE 82 AA; 8437 MW; 73A012103B4E9DF0 CRC64;

QY 1 LVIGVPAKIQF 12
 :||| | :|||

Db 41 LVIGVPEKTSF 52

RESULT 15

Q92R37 PRELIMINARY; PRT; 201 AA.

AC Q92R37; 01-DEC-2001 (TREMblrel. 19, Created)

DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)

DE 01-MAR-2002 (TREMblrel. 20, Last annotation update)

DE Hypothetical protein R01087.

GN R01087 OR SMC02636.

OS Rhizobium meliloti (Sinorhizobium meliloti).

OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;

OC Rhizobiaceae; Sinorhizobium.

NCBI_TaxID=382;

RM [1]

RP SEQUENCE FROM N.A.

RC STRAIN=1021;

RX MEDLINE=21396507; PubMed=11481430;

RA Capela D., Barloy-Hubler F., Gouzy J., Bothe G., Ampe F., Batut J.,

RA Boistard P., Becker A., Boutry M., Cadieu E., Dreano S., Gloux S.,

RA Godrie T., Goffeau A., Kahn D., Kiss E., Lelaure V., Masuy D.,

RA Pohl T., Portetelle D., Puehler A., Purnelle B., Ramsperger U.,

RA Renard G., Thebaud P., Vandenbol M., Weidner S., Galibert F.,

RT "Analysis of the chromosome sequence of the legume symbiont

RT Sinorhizobium meliloti strain 1021."
 RL Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).
 DR EMBL: AL591786; CAC45666.1; -;
 KW Hypothetical protein, complete proteome.
 DR SEQUENCE 201 AA; 21985 MW; 6610A10C3F5DDC84 CRC64;

Query Match 57.1%; Score 40; DB 16; Length 201;
 Best Local Similarity 58.3%; Pred. No. 33;
 Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 3 LGVGPAKIQFEG 14
 :||| | :|||
 Db 63 LGVGPAGLEFEG 74

Search completed: January 13, 2003, 09:59:42
 Job time : 56.1282 secs

Mon Jan 13 10:35:08 2003

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 13, 2003, 09:53:20 : Search time 12.2051 Seconds
(without alignments)
33.750 Million cell updates/sec

Title: US-09-554-941-1
Perfect score: 70
Sequence: 1 VILGVPKIOFEG 14

Scoring table: BIOSUM62
Gapop 10.0, Gapext 0.5

Searched: 262574 seqs, 29422922 residues
Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents, AA: *
1: /cgn2_6/ptodata/1/aa/5A.COMB.pep.*
2: /cgn2_6/ptodata/1/aa/5B.COMB.pep.*
3: /cgn2_6/ptodata/1/aa/6A.COMB.pep.*
4: /cgn2_6/ptodata/1/aa/6B.COMB.pep.*
5: /cgn2_6/ptodata/1/aa/PCTUS.COMB.pep.*
6: /cgn2_6/ptodata/1/aa/Backfilest.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	70	100.0	75	1 US-08-322-962-5	Sequence 5, Appl
2	70	100.0	75	3 US-08-450-653-5	Sequence 5, Appl
3	37	52.9	566	4 US-09-491-522-7	Sequence 7, Appl
4	37	52.9	1205	4 US-09-491-522-11	Sequence 11, Appl
5	37	52.9	1211	4 US-09-491-522-5	Sequence 5, Appl
6	36	51.4	194	4 US-08-968-683A-35	Sequence 35, Appl
7	36	51.4	194	4 US-08-968-683A-35	Sequence 4, Appl
8	36	51.4	414	4 US-08-914-375C-72	Sequence 2, Appl
9	36	51.4	459	4 US-08-063-035-2	Sequence 319, Appl
10	35	50.0	141	4 US-09-613-192A-319	Sequence 10, Appl
11	35	50.0	395	4 US-09-036-987A-10	Sequence 12, Appl
12	35	50.0	395	4 US-09-370-700-10	Sequence 10, Appl
13	35	50.0	516	4 US-08-948-564-12	Sequence 6, Appl
14	35	50.0	516	4 US-08-948-564-12	Sequence 12, Appl
15	35	50.0	686	2 US-08-993-228-12	Sequence 2, Appl
16	35	50.0	686	2 US-08-993-228-12	Sequence 458, Appl
17	35	50.0	737	4 US-08-936-165A-458	Sequence 5, Appl
18	34	48.6	74	2 US-08-785-795-5	Sequence 7, Appl
19	34	48.6	296	4 US-08-655-270A-7	Sequence 7, Appl
20	34	48.6	296	4 US-08-651-841-7	Sequence 7, Appl
21	34	48.6	296	4 US-09-955-597-7	Sequence 7, Appl
22	34	48.6	403	4 US-09-503-391-8	Sequence 14, Appl
23	34	48.6	403	4 US-07-754-918A-14	Sequence 14, Appl
24	34	48.6	766	4 US-09-134-001C-3689	Sequence 3689, Ap
25	34	48.6	767	4 US-09-134-001C-4023	Sequence 4023, Ap
26	34	48.6	1043	2 US-08-724-354D-4	Sequence 4, Appl
27	34	48.6	1043	2 US-08-724-354D-4	Sequence 4, Appl

28	34	48.6	1043	3 US-09-270-984A-4	Sequence 4, Appl
29	34	48.6	1118	2 US-08-724-354D-2	Sequence 2, Appl
30	34	48.6	1118	3 US-09-270-984A-2	Sequence 2, Appl
31	33.5	47.9	1958	1 US-07-945-283-2	Sequence 9, Appl
32	33	47.1	28	1 US-08-427-072-9	Sequence 37, Appl
33	33	47.1	311	1 US-08-118-270-37	Sequence 37, Appl
34	33	47.1	311	5 PCT-US93-08528-37	Sequence 4857, Ap
35	33	47.1	348	4 US-09-134-001C-4857	Sequence 4, Appl
36	33	47.1	361	1 US-08-685-945B-4	Sequence 4, Appl
37	33	47.1	363	4 US-08-685-945B-4	Sequence 44, Appl
38	33	47.1	365	1 US-09-108-020-44	Sequence 2, Appl
39	33	47.1	365	1 US-08-390-162-2	Sequence 2, Appl
40	33	47.1	365	1 US-08-685-945B-2	Sequence 8, Appl
41	33	47.1	365	3 US-08-155-005A-8	Sequence 8, Appl
42	33	47.1	365	3 US-09-363-783-8	Sequence 4, Appl
43	33	47.1	378	1 US-08-416-756A-4	Sequence 4, Appl
44	33	47.1	378	4 US-08-880-965-4	Sequence 3, Appl
45	33	47.1	385	1 US-08-416-756A-3	Sequence 3, Appl

ALIGNMENTS

RESULT 1
US-08-322-962-5
; Sequence 5, Application US/08322962
; Patent No. 5466785
; GENERAL INFORMATION:
; APPLICANT: Deftamond, Annick J
; TITLE OF INVENTION: Tissue-Preferential Promoters
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: CIBA-GEIGY Corporation
; STREET: 7 Skyline Drive
; CITY: Hawthorne
; STATE: New York
; COUNTRY: USA
; ZIP: 10532
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30B
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/322,962
; FILING DATE:
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/071,209
; FILING DATE: 02-JUN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/508,207
; FILING DATE: 12-APR-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: SPTULLI, W. Murray
; REGISTRATION NUMBER: 32,943
; REFERENCE/DOCKET NUMBER: S-18039/CGC 1479/CONT2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-541-8615
; TELEFAX: 919-541-8689
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 75 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: Protein
; LOCATION: 1..75
; OTHER INFORMATION: /note= "Protein product of pea
; OTHER INFORMATION: metallothionein reported by Evans et al."

Query Match
Best Local Similarity 100.0%; Score 70; DB 1; Length 75;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 VILGVGPAKIOFEG 14
DB 36 VILGVGPAKIOFEG 49

RESULT 2

US-08-450-653-5
; Sequence 5, Application US/08450653
; Patent No. 6018099
; GENERAL INFORMATION:
; APPLICANT: DeFranco, Annick J
; TITLE OF INVENTION: Tissue-Preferential Promoters
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CIBA-GEIGY Corporation
; STREET: 7 Skyline Drive
; CITY: Hawthorne
; STATE: New York
; COUNTRY: USA
; ZIP: 10532
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30B
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/450,653
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/322,962
; FILING DATE:
; INFORMATION FOR SEQ ID NO: 1
; ATTORNEY/AGENT INFORMATION:
; NAME: Spullin, W. Murray
; REGISTRATION NUMBER: 32,943
; REFERENCE/DOCKET NUMBER: S-18039/CGC 1479/CONT2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-541-8615
; TELEFAX: 919-541-8689
; INFORMATION FOR SEQ ID NO: 5
; SEQUENCE CHARACTERISTICS:
; LENGTH: 75 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: Protein
; LOCATION: 1..75
; OTHER INFORMATION:
; OTHER INFORMATION: /note="Protein product of pea
; US-08-450-653-5 metallothionein reported by Evans et al."

Query Match
Best Local Similarity 100.0%; Score 70; DB 3; Length 75;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 VILGVGPAKIOFEG 14
DB 36 VILGVGPAKIOFEG 49

RESULT 3
US-09-491-522-7
; Sequence 7, Application US/09491522

Patent No. 6428998
; GENERAL INFORMATION:
; APPLICANT: Collige, Alain
; APPLICANT: Lapierre, Charles M.
; APPLICANT: Prockop, Darwin J.
; TITLE OF INVENTION: RECOMBINANT N-PROTEINASE,
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds, LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036-2811
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: Windows
; SOFTWARE: FASTSEQ for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/491,522
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/886,333
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Abrams, Samuel B
; REGISTRATION NUMBER: 30,605
; REFERENCE/DOCKET NUMBER: 8389-0060-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-493-4935
; TELEFAX: 650-493-5556
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 7
; SEQUENCE CHARACTERISTICS:
; LENGTH: 566 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-09-491-522-7
Query Match
Best Local Similarity 52.9%; Score 37; DB 4; Length 566;
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
QY 1 VILGVGPAKIOFEG 14
DB 270 VILGVDSVVOFG 283

RESULT 4
US-09-491-522-11
; Sequence 11, Application US/09491522
; Patent No. 6428998
; GENERAL INFORMATION:
; APPLICANT: Collige, Alain
; APPLICANT: Lapierre, Charles M.
; APPLICANT: Prockop, Darwin J.
; TITLE OF INVENTION: RECOMBINANT N-PROTEINASE,
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds, LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036-2811
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible

Mon Jan 13 10:35:08 2003

us-09-554-941-1.rai

Page 3

OPERATING SYSTEM: Windows
SOFTWARE: FASTSEQ for Windows Version 2.0b
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/491,522
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA: 08/886,333
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Abrams, Samuel B
REGISTRATION NUMBER: 30,605
REFERENCE/DOCKET NUMBER: 8389-0060-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-493-4935
TELEFAX: 650-493-5556
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 1205 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-09-491-522-11
Query Match 52.9%; Score 37; DB 4; Length 1205;
Best local Similarity 50.0%; Pred. No. 1.8e+02;
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
OY 1 VILGVGPAKIOFEG 14
DB 264 VILGVDDSVVOFEG 277
RESULT 5
US-09-491-522-5
Sequence 5, Application US/09491522
Patent No. 642898
GENERAL INFORMATION:
APPLICANT: Collige, Alain
APPLICANT: Lapierre, Charles M.
APPLICANT: Prockop, Darin J.
TITLE OF INVENTION: RECOMBINANT N-PROTEINASE,
TITLE OF INVENTION: AND THE PRODUCTION, METHODS AND USES THEREOF
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds, LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10036-2811
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows
SOFTWARE: FASTSEQ for Windows Version 2.0b
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/491,522
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/886,333
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Abrams, Samuel B
REGISTRATION NUMBER: 30,605
REFERENCE/DOCKET NUMBER: 8389-0060-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-493-4935
TELEFAX: 650-493-5556
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:
LENGTH: 1211 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-09-491-522-5
Query Match 52.9%; Score 37; DB 4; Length 1211;
Best local Similarity 50.0%; Pred. No. 1.8e+02;
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
OY 1 VILGVGPAKIOFEG 14
DB 270 VILGVDDSVVOFEG 283
RESULT 6
US-08-968-563-35
Sequence 35, Application US/08968563
Patent No. 6013494
GENERAL INFORMATION:
APPLICANT: CHARLES E. NAKAMURA
APPLICANT: ANTHONY A. GATENBY
APPLICANT: AMY (KUANG-HUI) HSU
APPLICANT: RICHARD D. LA REAU
APPLICANT: SHARON L. HAYNIE
APPLICANT: MARIA DIAZ-TORRES
APPLICANT: DONALD E. TRIMBUR
APPLICANT: GREGORY M. WHITED
APPLICANT: VASANTHA NAGARAJAN
APPLICANT: MARK S. PAYNE
APPLICANT: STEPHEN K. PICATAGGIO
APPLICANT: RAMESCH V. NAIR
TITLE OF INVENTION: METHOD FOR THE RECOMBINANT
TITLE OF INVENTION: PRODUCTION OF 1,3-PROPANDIOL
NUMBER OF SEQUENCES: 49
CORRESPONDENCE ADDRESS:
ADDRESSEE: E. I. DU PONT DE NEMOURS AND COMPANY
STREET: 1007 MARKET STREET
CITY: WILMINGTON
STATE: DELAWARE
COUNTRY: U.S.A.
ZIP: 19898
ADDRESSEE: GENENCOR INTERNATIONAL, INC.
STREET: 4 CAMBRIDGE PLACE
STREET: 1870 SOUTH WINTON ROAD
CITY: ROCHESTER
STATE: NEW YORK
COUNTRY: U.S.A.
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.50 INCH DISKETTE
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: MICROSOFT WORD FOR WINDOWS 95
SOFTWARE: MICROSOFT WORD VERSION 7.0A
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/968,563
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA: 60/030,601
APPLICATION NUMBER: NOVEMBER 13, 1996
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: FLOYD, LINDA AXAMETHY
REGISTRATION NUMBER: 33,692
REFERENCE/DOCKET NUMBER: CR-9982
TELECOMMUNICATION INFORMATION:
TELEPHONE: 302-773-0164
TELEFAX: 302-773-0164
INFORMATION FOR SEQ ID NO: 35:
SEQUENCE CHARACTERISTICS:
LENGTH: 194 amino acids
TYPE: amino acid
STRANDEDNESS: unknown

TOPOLOGY: unknown
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: DHAB2
US-08-968-563-35

Query Match
Best Local Similarity 51.4%; Score 36; DB 3; Length 194;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
OY 1 VIIIGVPA 8
Db 30 VVIGVPA 37

RESULT 7
US-08-969-683A-35
Sequence 35, Application US/08969683A
Patent No. 6136576
GENERAL INFORMATION:
APPLICANT: GENENCOR INTERNATIONAL, INC.
TITLE OF INVENTION: METHOD FOR THE RECOMBINANT
NUMBER OF SEQUENCES: 68
CORRESPONDENCE ADDRESS:
ADDRESS: Genencor International, Inc.
STREET: 4 Cambridge Place
CITY: Rochester
STATE: NY
COUNTRY: U.S.A
ZIP: 14618
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM compatible
OPERATING SYSTEM: Windows
SOFTWARE: FASTSEQ for Windows Version 2.0b
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/969,683A
FILING DATE:
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: PCT/US97/20873
FILING DATE: 13-NOV-1997
APPLICATION NUMBER: 60/030,601
FILING DATE: 13-NOV-1996
ATTORNEY/AGENT INFORMATION:
NAME: Glaister, Dedra
REGISTRATION NUMBER: 33,888
REFERENCE/DOCKET NUMBER: GC 369-2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-864-7620
TELEFAX: 650-845-6504
TELEX:
INFORMATION FOR SEQ ID NO: 35:
SEQUENCE CHARACTERISTICS:
LENGTH: 194 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: DHAB2
US-08-969-683A-35

Query Match
Best Local Similarity 51.4%; Score 36; DB 4; Length 194;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
OY 1 VIIIGVPA 8
Db 30 VVIGVPA 37

RESULT 8
US-09-608-917A-4
Sequence 4, Application US/09608917A
Patent No. 6409648
GENERAL INFORMATION:
APPLICANT: The Regents of the University of California
APPLICANT: Campisi, Judith
TITLE OF INVENTION: TRF1 Binding Protein
FILE REFERENCE: LBNL IB 1317
CURRENT APPLICATION NUMBER: US/09/608,917A
NUMBER OF SEQ ID NOS: 4
SOFTWARE: Patent in version 3.0
SEQ ID NO 4
LENGTH: 414
TYPE: PRT
ORGANISM: mouse
US-09-608-917A-4

Query Match
Best Local Similarity 51.4%; Score 36; DB 4; Length 414;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
OY 4 GVGPAKIQ 12
Db 6 GVGPAKLR 14

RESULT 9
US-08-914-375C-72
Sequence 72, Application US/08914375C
Patent No. 6377893
GENERAL INFORMATION:
APPLICANT: Steven A. Benner
NUMBER OF SEQUENCES: 74
CORRESPONDENCE ADDRESS:
ADDRESS: Steven A. Benner
STREET: 1501 NW 68th Terrace
CITY: Gainesville
STATE: FL
COUNTRY: United States
ZIP: 32605-4147
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch diskette
COMPUTER: Apple Macintosh
OPERATING SYSTEM: Macintosh 7.0
SOFTWARE: Microsoft Word
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/914,375C
FILING DATE: 19-Aug-1997
CLASSIFICATION: 702/20
TELECOMMUNICATION INFORMATION:
TELEPHONE: 352 392 7773
TELEFAX: 352 331 0462
INFORMATION FOR SEQ ID NO: 72:
SEQUENCE CHARACTERISTICS:
LENGTH: 459
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: linear
ORIGINAL SOURCE:
ORGANISM: Trifolium repens
FEATURE:
OTHER INFORMATION: bgl2-trip non-cyanogenic strand-glucosidase precu
US-08-914-375C-72

Query Match
Best Local Similarity 51.4%; Score 36; DB 4; Length 459;
Matches 7; Conservative 1; Mismatches 5; Indels 0; Gaps 0;


```

QY      2  ILGVGPAKIQFEFG  14
          | | | : | | |
Db      11 IFGAGSSAYQFEFG  23

```

GENERAL INFORMATION:
APPLICANT: PEUKERT, Karen; HAENEL, Frank; and EILERS,
APPLICANT: Matlin
TITLE OF INVENTION: Myc-binding zinc finger proteins,
TITLE OF INVENTION: Their preparation and their use
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Keil & Weinkauff
STREET: 1101 Connecticut Avenue
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20036
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 1.2 Mb storage
COMPUTER: IBM AT-compatible, 80486 processor
OPERATING SYSTEM: MS-DOS version 6.1
SOFTWARE: WordPerfect version 8.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/063,035
FILING DATE: 21-APR-1998
CLASSIFICATION: 514
INFORMATION FOR SEQ. ID NO.: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 803 amino acids
TYPE: amino acid
TOPOLOGY: linear

US-09-063-035-2

Query Match 51.4%; Score 36; DB 4; Length 803;
Best Local Similarity 54.5%;
Matches 6; Conservative 3; Mismatches 2; Indels

ESTIMUT 11
 US-09-615-192A-319
 Sequence 319, Application US/09615192A
 Patent No. 6410718
 GENERAL INFORMATION:
 APPLICANT: Bloksberg, Leonard N.
 APPLICANT: Havukari, Ilkka
 TITLE OF INVENTION: Materials and Methods for the
 TITLE OF INVENTION: Modification of Plant Lignin Content
 FILE REFERENCE: 11000.1003c4U
 CURRENT APPLICATION NUMBER: US/09/615,192A
 CURRENT FILING DATE: 2000-07-12
 PRIOR APPLICATION NUMBER: US 08/975,316
 PRIOR FILING DATE: 1997-11-21
 PRIOR APPLICATION NUMBER: US 08/713,000
 PRIOR FILING DATE: 1996-09-11
 PRIOR APPLICATION NUMBER: US 09/169,789
 PRIOR FILING DATE: 1998-10-09
 NUMBER OF SEQ ID NOS: 405
 SOFTWARE: FASTSEQ for Windows Version 3.0
 SEQ ID NO 319
 LENGTH: 141
 TYPE: PRT
 ORGANISM: Eucalyptus grandis
 US-09-615-192A-319

```
QY      1 VILGVGPAKIQF 12
         |||| | ||: |
Db      71 VILGDNKAVVF 82
```

GENERAL INFORMATION:
APPLICANT: Baltz, Richard H.
APPLICANT: Broughton, Mary C.
APPLICANT: Crawford, Kathryn P.
APPLICANT: Madduri, Krishnamurthy
APPLICANT: Merlo, Donald J.
APPLICANT: Treadway, Patli J.
APPLICANT: Turner, Jan R.
APPLICANT: Waldron, Clive
TITLE OF INVENTION: Biosynthetic Genes For Spinosyn Insecticide
TITLE OF INVENTION: Production
NUMBER OF SEQUENCES: 39
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dow Agrosciences LLC Patent Department
STREET: 9330 Zionsville Road
CITY: Indianapolis
STATE: Indiana
COUNTRY: USA
ZIP: 46268
COMPUTER READERABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/036,987A
FILING DATE: 09-MAR-1998
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Stuart, Donald R.
REGISTRATION NUMBER: 28,479
REFERENCE/DOCKET NUMBER: 50,608
TELECOMMUNICATION INFORMATION:
TELEPHONE: (317)337-4816
TELEFAX: (317)337-4847
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 395 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-036-987A-10

RESULT 13
US-09-370-700-10
; Sequence 10, Application US/09370700
; Patent No. 6274350
; GENERAL INFORMATION:
; APPLICANT: Baltz, Richard H
; APPLICANT: Broughton, Mary C

APPLICANT: Crawford, Kathryn P
APPLICANT: Madduri, Krishnamurthy
APPLICANT: Treadway, Patli J
APPLICANT: Turner, Jan R
TITLE OF INVENTION: Biosynthetic Genes For Spinosyn Insecticide
FILE REFERENCE: 50489 DIV1
CURRENT APPLICATION NUMBER: US/09/370,700
CURRENT FILING DATE: 1999-08-09
EARLIER APPLICATION NUMBER: US 09/36987
NUMBER OF SEQ ID NOS: 39
SOFTWARE: Patentln Ver. 2.0
SEQ ID NO: 10
LENGTH: 395
TYPE: PRT
ORGANISM: Saccharopolyspora spinosa
US-09-370-700-10

Query Match
Best Local Similarity 50.0%; Score 35; DB 4; Length 395;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 4 GVGPAKIQ 11
Db 20 GAGPAKIQ 27

RESULT 14
US-08-948-564-12
Sequence 12, Application US/08948564
Patent No. 6121512
GENERAL INFORMATION:
APPLICANT: Siminszky, Balazs
APPLICANT: Dewey, Ralph E.
TITLE OF INVENTION: Cordin, Frederick T.
TITLE OF INVENTION: No. 6121512el Cytochrome P-450 Constructs and
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESS: Virginia C. Bennett
STREET: PO Box 37428
CITY: Raleigh
STATE: No. 6121512elh Carolina
COUNTRY: USA
ZIP: 27627
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/948,564
FILING DATE:
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Bennett, Virginia C.
REGISTRATION NUMBER: 37,092
REFERENCE/DOCKET NUMBER: 5051-409
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-854-1400
TELEFAX: 919-854-1401
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 516 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-948-564-12

Query Match
Best Local Similarity 50.0%; Score 35; DB 3; Length 516;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 3 LGVGPAKIOFEG 14
Db 99 LFGVPAKLSYDG 110

RESULT 15
US-09-201-641-6
Sequence 6, Application US/09201641A
Patent No. 6232530
GENERAL INFORMATION:
APPLICANT: Cunningham Jr, Francis X
APPLICANT: Dellapenna, Dean
TITLE OF INVENTION: Method for Regulating Carotenoid Biosynthesis in
FILE REFERENCE: Quest 41-162
CURRENT APPLICATION NUMBER: US/09/201,641A
CURRENT FILING DATE: 1998-11-30
NUMBER OF SEQ ID NOS: 8
SOFTWARE: Patentln Ver. 2.0
SEQ ID NO: 6
LENGTH: 516
TYPE: PRT
ORGANISM: Tagetes erecta
US-09-201-641-6

Query Match
Best Local Similarity 50.0%; Score 35; DB 4; Length 516;
Matches 6; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

OY 1 VILGPAKIOFEG 14
Db 98 VVIGGPAKGLAG 111

Search completed: January 13, 2003, 09:56:22
Job time : 14.2051 secs

Title:	US-09-554-941-1
Perfect score:	70
Sequence:	1 VILGVGPAKIQFEG 14

Searched: 118974 seqs, 19401057 residues
Chosen parameters: 11897

```

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing:      Minimum Match 0%
                      Maximum Match 100%
                      Listing first 45 summaries

```

```

1: published_applications_AA:*
2: /cgn2_6/ptodata/1/pubppaa/US08_NEW_PUB pep:*
3: /cgn2_6/ptodata/1/pubppaa/PCTO_NEW_PUB pep:*
4: /cgn2_6/ptodata/1/pubppaa/US06_NEW_PUB pep:*
5: /cgn2_6/ptodata/1/pubppaa/US07_PUBCOMB pep:*
6: /cgn2_6/ptodata/1/pubppaa/US07_PUBCOMB pep:*
7: /cgn2_6/ptodata/1/pubppaa/PCTO5_PUBCOMB pep:*
8: /cgn2_6/ptodata/1/pubppaa/US08_PUBCOMB pep:*
9: /cgn2_6/ptodata/1/pubppaa/US09_NEW_PUB pep:*
10: /cgn2_6/ptodata/1/pubppaa/US09_PUBCOMB pep:*
11: /cgn2_6/ptodata/1/pubppaa/US10_NEW_PUB pep:*
12: /cgn2_6/ptodata/1/pubppaa/US10_PUBCOMB pep:*
13: /cgn2_6/ptodata/1/pubppaa/US06_NEW_PUB pep:*
14: /cgn2_6/ptodata/1/pubppaa/US06_PUBCOMB pep:*

```

14: /cguz2-w-f-
Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed.
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	% Length	DB	ID	Description
1	39	55.7	28	10	US-09-864-761-66875	Sequence 36875, A
2	37	52.9	969	10	US-09-921-98783-5	Sequence 5, App
3	35	50.0	366	10	US-09-801-368-254	Sequence 254, App
4	35	50.0	457	9	US-09-738-626-6965	Sequence 6965, App
5	35	50.0	510	9	US-09-738-626-5777	Sequence 5777, App
6	35	50.0	510	9	US-09-323-998D-54	Sequence 2, App1
7	35	50.0	516	10	US-09-889-882-2	Sequence 458, App
8	35	50.0	737	10	US-09-999-980-458	Sequence 30, App1
9	34	48.6	70	10	US-09-764-884-10	Sequence 279, App
10	34	48.6	121	9	US-09-766-887-279	Sequence 203, App
11	34	48.6	125	10	US-09-766-887-203	Sequence 203, App
12	34	48.6	135	10	US-09-922-217-203	Sequence 203, App
13	34	48.6	135	10	US-09-833-263-203	Sequence 45, App1
14	34	48.6	359	9	US-09-978-295A-45	Sequence 45, App1
15	34	48.6	359	9	US-09-978-697-45	Sequence 45, App1
16	34	48.6	359	9	US-09-978-192K-45	Sequence 45, App1
17	34	48.6	359	9	US-09-999-832K-45	Sequence 45, App1
18	34	48.6	359	9	US-09-978-189-45	Sequence 45, App1
19	34	48.6	359	12	US-10-052-586-42	Sequence 42, App1

[illegible]

ALIGNMENTS

RESULT 1
 US-09-864-761-36875
 Sequence 36875, Application US/09864761
 Patent No. US20020048763A1
 GENERAL INFORMATION:
 APPLICANT: Penn, Sharon G.
 APPLICANT: Rank, David R.
 APPLICANT: Hanzel, David K.
 APPLICANT: Chen, Wensheng
 TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
 TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
 FILE REFERENCE: Aesomica-X-1
 CURRENT APPLICATION NUMBER: US/09/864,761
 CURRENT FILING DATE: 2001-05-23
 PRIOR APPLICATION NUMBER: US 60/180,312
 PRIOR FILING DATE: 2000-02-04
 PRIOR APPLICATION NUMBER: US 60/207,456
 PRIOR FILING DATE: 2000-05-26
 PRIOR APPLICATION NUMBER: US 09/632,366
 PRIOR FILING DATE: 2000-08-03
 PRIOR APPLICATION NUMBER: GB 24263.6
 PRIOR FILING DATE: 2000-10-04
 PRIOR APPLICATION NUMBER: US 60/236,359
 PRIOR FILING DATE: 2000-09-27
 PRIOR APPLICATION NUMBER: PCT/US01/00666
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00667
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00664
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00669
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00665
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00668
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00663
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00662
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00661
 PRIOR FILING DATE: 2001-01-30

```

; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 36875
; LENGTH: 28
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL049539.19
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 2.3
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.9
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.4
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.5
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.6
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 3.3
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.3
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.2
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.4
; OTHER INFORMATION: EST_HUMAN HIT: AV659818.1, EVALUATE 4.00e-08
; US-09-864-761-36875

```

```

Query Match
Best Local Similarity 55.7%; Score 39; DB 10; Length 28;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

```

```

Qy 2 ILGVGPAKIQFE 13
Db 7 VLAAGPQLOAE 18

```

```

RESULT 2
; Sequence 5, Application US/09321987B
; Patent No. US20020102210A1
; GENERAL INFORMATION:
; APPLICANT: Kimble, Judith E
; TITLE OF INVENTION: Agent and Method for Modulating Cell Migration
; FILE REFERENCE: 960296.95386
; CURRENT APPLICATION NUMBER: US/09/321,987B
; PRIOR FILING DATE: 1999-05-28
; PRIOR APPLICATION NUMBER: 60/087,170
; PRIOR FILING DATE: 1998-05-29
; PRIOR APPLICATION NUMBER: 60/129,023
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 969
; TYPE: PRT
; ORGANISM: Bovine
; US-09-321-987B-5

```

```

Query Match
Best Local Similarity 52.9%; Score 37; DB 10; Length 969;
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
Qy 1 VILGVGPAKIQFE 14
Db 264 VILGVDSVQFHHG 277

```

```

RESULT 3
US-09-801-368-254
; Sequence 254, Application US/09801368

```

```

; Patent No. US20020128250A1
; GENERAL INFORMATION:
; APPLICANT: Busby, Robert
; APPLICANT: Call, Brian
; APPLICANT: Hecht, Peter
; APPLICANT: Holtzman, Doug
; APPLICANT: Madden, Kevin
; APPLICANT: Maxon, Mary
; APPLICANT: Milne, Todd
; APPLICANT: No. US20020128250A1man, Thea
; APPLICANT: Royer, John
; APPLICANT: Salama, Sofie
; APPLICANT: Sherman, Amir
; APPLICANT: Silva, Jeff
; TITLE OF INVENTION: Methods for Improving Secondary Metabolite Production in Fungi
; FILE REFERENCE: 109272.147
; CURRENT APPLICATION NUMBER: US/09/801,368
; PRIOR FILING DATE: 2001-03-07
; PRIOR APPLICATION NUMBER: US 09/487,558
; PRIOR FILING DATE: 2000-01-19
; PRIOR APPLICATION NUMBER: US 60/160,587
; NUMBER OF SEQ ID NOS: 440
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 254
; LENGTH: 366
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
; US-09-801-368-254

```

```

Query Match
Best Local Similarity 50.0%; Score 35; DB 10; Length 366;
Matches 5; Conservative 5; Mismatches 3; Indels 0; Gaps 0;
Qy 2 ILGVGPAKIQFE 14
Db 116 VMAVMPPEVQFDG 128

```

```

RESULT 4
; Sequence 6965, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, MAKINO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 6965
; LENGTH: 467
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
; US-09-738-626-6965

```

APPLICANT: Black, Michael
 Burnham, Martin
 Hodgson, John
 Knowles, David
 Lonetto, Michael
 Nicholas, Richard
 Pratt, Julie
 Reichard, Richard
 Rosenberg, Martin
 Ward, Judith
 TITLE OF INVENTION: No. US20020082234A1el Prokaryotic Polynucleotides
 Polypeptides and Their Uses
 NUMBER OF SEQUENCES: 534
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Smithline Beecham Corporation
 STREET: 709 Swedeland Road
 CITY: King of Prussia
 STATE: PA
 COUNTRY: USA
 ZIP: 19406-0939
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette

```

; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/939,980
; FILING DATE: 27-Aug-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/936,165
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Gimm1, Edward R
; REGISTRATION NUMBER: 38,891
; REFERENCE/DOCKET NUMBER: P50549
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-4478
; TELEFAX: 610-270-5090
; TELEX: <Unknown>
; INFORMATION FOR SEQ ID NO: 458:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 70 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 458:
US-09-939-980-458

Query Match      48.6% Score 34; DB 10; Length 70;
Best Local Similarity 70.0%; Pred. No. 13;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 GVGPAKIQFE 13
   ||| | |
Db 42 GVGPKENKE 51

RESULT 9
US-09-764-884-30
; Sequence 30, Application US/09764884
; Patent No. US20020161208A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PRT03
; CURRENT APPLICATION NUMBER: US/09/764,884
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 30
; LENGTH: 121
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-764-884-30

Query Match      48.6% Score 34; DB 9; Length 121;
Best Local Similarity 38.5%; Pred. No. 26;
Matches 5; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 2 ILGVGPAKIQFEG 14
   :|:| | |
Db 96 VIGIAGKIVFDS 108

RESULT 10
US-09-764-887-279
; Sequence 279, Application US/09764887
; Patent No. US20020042096A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PAl13

```

```

; CURRENT APPLICATION NUMBER: US/09/764,887
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 658
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 279
; LENGTH: 125
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-764-887-279

Query Match      48.6% Score 34; DB 10; Length 125;
Best Local Similarity 50.0%; Pred. No. 27;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 3 LGVGPAKIQFEG 14
   ||| | |
Db 92 LGMGPKAEPQG 103

RESULT 11
US-10-025-380-203
; Sequence 203, Application US/10025380
; Publication No. US20020182191A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Lodes, Michael J.
; APPLICANT: Secrist, Heather
; APPLICANT: Benson, Darin R.
; APPLICANT: Meagher, Madeleine Joy
; APPLICANT: Stoik, John A.
; APPLICANT: Wang, Tonglong
; APPLICANT: Jiang, Yugu
; APPLICANT: Smith, Carole L.
; APPLICANT: King, Gordon E.
; APPLICANT: Wang, Aljun
; APPLICANT: Clapper, Jonathan D.
; APPLICANT: Skelky, Yasir A. W.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS
; FILE REFERENCE: 210121.471C14
; CURRENT APPLICATION NUMBER: US/10/025,380
; CURRENT FILING DATE: 2001-12-19
; NUMBER OF SEQ ID NOS: 1129
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 203
; LENGTH: 135
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-025-380-203

Query Match      48.6% Score 34; DB 9; Length 135;
Best Local Similarity 54.5%; Pred. No. 29;
Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 4 GVGPAKIQFEG 14
   ||| | |
Db 80 GVNPGKIDVEG 90

RESULT 12
US-09-922-217-203
; Sequence 203, Application US/09922217
; Patent No. US20020076414A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Lodes, Michael J.
; APPLICANT: Secrist, Heather
; APPLICANT: Benson, Darin R.
; APPLICANT: Meagher, Madeleine Joy

```

APPLICANT: Gafo, Wei-Qiang
 APPLICANT: Gerber, Hanspeter
 APPLICANT: Gerritsen, Mary E.
 APPLICANT: Goddard, Audrey
 APPLICANT: Godowski, Paul J.
 APPLICANT: Grimaldi, J. Christopher
 APPLICANT: Gurney, Austin L.
 APPLICANT: Hillan, Kenneth J.
 APPLICANT: Kijavlin, Ivar J.
 APPLICANT: Kuio, Sophia S.
 APPLICANT: Napier, Mary A.
 APPLICANT: Pan, James
 APPLICANT: Paoni, Nicholas F.
 APPLICANT: Roy, Margaret Ann
 APPLICANT: Shelton, David L.
 APPLICANT: Stewart, Timothy A.
 APPLICANT: Tumas, Daniel
 APPLICANT: Williams, P. Mickey
 APPLICANT: Wood, William I.
 TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
 TITLE OF INVENTION: Acids Encoding the Same
 FILE REFERENCE: P2630P1C11
 CURRENT APPLICATION NUMBER: US/09/978,295A
 CURRENT FILING DATE: 2001-10-15
 PRIOR APPLICATION NUMBER: 09/918585
 PRIOR FILING DATE: 2001-07-30
 PRIOR APPLICATION NUMBER: 60/062250
 PRIOR FILING DATE: 1997-10-17
 PRIOR APPLICATION NUMBER: 60/064249
 PRIOR FILING DATE: 1997-11-03
 PRIOR APPLICATION NUMBER: 60/065311
 PRIOR FILING DATE: 1997-11-13
 PRIOR APPLICATION NUMBER: 60/066364
 PRIOR FILING DATE: 1997-11-21
 PRIOR APPLICATION NUMBER: 60/077450
 PRIOR FILING DATE: 1998-03-10
 PRIOR APPLICATION NUMBER: 60/077632
 PRIOR FILING DATE: 1998-03-11
 PRIOR APPLICATION NUMBER: 60/077641
 PRIOR FILING DATE: 1998-03-11
 PRIOR APPLICATION NUMBER: 60/077649
 PRIOR FILING DATE: 1998-03-11
 PRIOR APPLICATION NUMBER: 60/077791
 PRIOR FILING DATE: 1998-03-12
 PRIOR APPLICATION NUMBER: 60/078004
 PRIOR FILING DATE: 1998-03-13
 PRIOR APPLICATION NUMBER: 60/078886
 PRIOR FILING DATE: 1998-03-20
 PRIOR APPLICATION NUMBER: 60/078936
 PRIOR FILING DATE: 1998-03-20
 PRIOR APPLICATION NUMBER: 60/078910
 PRIOR FILING DATE: 1998-03-20
 PRIOR APPLICATION NUMBER: 60/078939
 PRIOR FILING DATE: 1998-03-20
 PRIOR APPLICATION NUMBER: 60/079294
 PRIOR FILING DATE: 1998-03-25
 PRIOR APPLICATION NUMBER: 60/079656
 PRIOR FILING DATE: 1998-03-26
 PRIOR APPLICATION NUMBER: 60/079664
 PRIOR FILING DATE: 1998-03-27
 PRIOR APPLICATION NUMBER: 60/079689
 PRIOR FILING DATE: 1998-03-27
 PRIOR APPLICATION NUMBER: 60/079663
 PRIOR FILING DATE: 1998-03-27
 PRIOR APPLICATION NUMBER: 60/079728
 PRIOR FILING DATE: 1998-03-27
 PRIOR APPLICATION NUMBER: 60/079786
 PRIOR FILING DATE: 1998-03-27
 PRIOR APPLICATION NUMBER: 60/079920
 PRIOR FILING DATE: 1998-03-30
 PRIOR APPLICATION NUMBER: 60/079923
 PRIOR FILING DATE: 1998-03-30
 PRIOR APPLICATION NUMBER: 60/080105

PRIOR FILING DATE: 1998-03-31
 PRIOR APPLICATION NUMBER: 60/080107
 PRIOR FILING DATE: 1998-03-31
 PRIOR APPLICATION NUMBER: 60/080165
 PRIOR FILING DATE: 1998-03-31
 PRIOR APPLICATION NUMBER: 60/080194
 PRIOR FILING DATE: 1998-03-31
 PRIOR APPLICATION NUMBER: 60/080327
 PRIOR FILING DATE: 1998-04-01
 PRIOR APPLICATION NUMBER: 60/080328
 PRIOR FILING DATE: 1998-04-01
 PRIOR APPLICATION NUMBER: 60/080333
 PRIOR FILING DATE: 1998-04-01
 PRIOR APPLICATION NUMBER: 60/080334
 PRIOR FILING DATE: 1998-04-01
 PRIOR APPLICATION NUMBER: 60/081070
 PRIOR FILING DATE: 1998-04-08
 PRIOR APPLICATION NUMBER: 60/081049
 PRIOR FILING DATE: 1998-04-08
 PRIOR APPLICATION NUMBER: 60/081071
 PRIOR FILING DATE: 1998-04-08
 PRIOR APPLICATION NUMBER: 60/081195
 PRIOR FILING DATE: 1998-04-08
 PRIOR APPLICATION NUMBER: 60/081203
 PRIOR FILING DATE: 1998-04-09
 PRIOR APPLICATION NUMBER: 60/081229
 PRIOR FILING DATE: 1998-04-09
 PRIOR APPLICATION NUMBER: 60/081955
 PRIOR FILING DATE: 1998-04-15
 PRIOR APPLICATION NUMBER: 60/081817
 PRIOR FILING DATE: 1998-04-15
 PRIOR APPLICATION NUMBER: 60/081819
 PRIOR FILING DATE: 1998-04-15
 PRIOR APPLICATION NUMBER: 60/081952
 PRIOR FILING DATE: 1998-04-15
 PRIOR APPLICATION NUMBER: 60/081838
 PRIOR FILING DATE: 1998-04-15
 PRIOR APPLICATION NUMBER: 60/082568
 PRIOR FILING DATE: 1998-04-21
 PRIOR APPLICATION NUMBER: 60/082569
 PRIOR FILING DATE: 1998-04-21
 PRIOR APPLICATION NUMBER: 60/082704
 PRIOR FILING DATE: 1998-04-22
 PRIOR APPLICATION NUMBER: 60/082804
 PRIOR FILING DATE: 1998-04-22
 PRIOR APPLICATION NUMBER: 60/082700
 PRIOR FILING DATE: 1998-04-22
 PRIOR APPLICATION NUMBER: 60/082797
 PRIOR FILING DATE: 1998-04-22
 PRIOR APPLICATION NUMBER: 60/082796
 PRIOR FILING DATE: 1998-04-23
 PRIOR APPLICATION NUMBER: 60/083336
 PRIOR FILING DATE: 1998-04-27
 PRIOR APPLICATION NUMBER: 60/083322
 PRIOR FILING DATE: 1998-04-28
 PRIOR APPLICATION NUMBER: 60/083392
 PRIOR FILING DATE: 1998-04-29
 PRIOR APPLICATION NUMBER: 60/083495
 PRIOR FILING DATE: 1998-04-29
 PRIOR APPLICATION NUMBER: 60/083496
 PRIOR FILING DATE: 1998-04-29
 PRIOR APPLICATION NUMBER: 60/083499
 PRIOR FILING DATE: 1998-04-29
 PRIOR APPLICATION NUMBER: 60/083545
 PRIOR FILING DATE: 1998-04-29
 PRIOR APPLICATION NUMBER: 60/083554
 PRIOR FILING DATE: 1998-04-29
 PRIOR APPLICATION NUMBER: 60/083558
 PRIOR FILING DATE: 1998-04-29
 PRIOR APPLICATION NUMBER: 60/083559
 PRIOR FILING DATE: 1998-04-29
 PRIOR APPLICATION NUMBER: 60/083500
 PRIOR FILING DATE: 1998-04-29

PRIOR APPLICATION NUMBER: 60/083742
 PRIOR FILING DATE: 1998-04-30
 PRIOR APPLICATION NUMBER: 60/084366
 PRIOR FILING DATE: 1998-05-05
 PRIOR APPLICATION NUMBER: 60/084414
 PRIOR FILING DATE: 1998-05-06
 PRIOR APPLICATION NUMBER: 60/084441
 PRIOR FILING DATE: 1998-05-06
 PRIOR APPLICATION NUMBER: 60/084637
 PRIOR FILING DATE: 1998-05-07
 PRIOR APPLICATION NUMBER: 60/084639
 PRIOR FILING DATE: 1998-05-07
 PRIOR APPLICATION NUMBER: 60/084640
 PRIOR FILING DATE: 1998-05-07
 PRIOR APPLICATION NUMBER: 60/084598
 PRIOR FILING DATE: 1998-05-07
 PRIOR APPLICATION NUMBER: 60/084600
 PRIOR FILING DATE: 1998-05-07
 PRIOR APPLICATION NUMBER: 60/084627
 PRIOR FILING DATE: 1998-05-07
 PRIOR APPLICATION NUMBER: 60/084643
 PRIOR FILING DATE: 1998-05-07
 PRIOR APPLICATION NUMBER: 60/085339
 PRIOR FILING DATE: 1998-05-13
 PRIOR APPLICATION NUMBER: 60/085338
 PRIOR FILING DATE: 1998-05-13
 PRIOR APPLICATION NUMBER: 60/085323
 PRIOR FILING DATE: 1998-05-13
 PRIOR APPLICATION NUMBER: 60/085582
 PRIOR FILING DATE: 1998-05-15
 PRIOR APPLICATION NUMBER: 60/085700
 PRIOR FILING DATE: 1998-05-15
 PRIOR APPLICATION NUMBER: 60/085689
 PRIOR FILING DATE: 1998-05-15
 PRIOR APPLICATION NUMBER: 60/085579
 PRIOR FILING DATE: 1998-05-15
 PRIOR APPLICATION NUMBER: 60/085580
 PRIOR FILING DATE: 1998-05-15
 PRIOR APPLICATION NUMBER: 60/085573
 PRIOR FILING DATE: 1998-05-15
 PRIOR APPLICATION NUMBER: 60/085704
 PRIOR FILING DATE: 1998-05-15
 PRIOR APPLICATION NUMBER: 60/085697

Query Match 48.6%; Score 34; DB 9; Length 359;
 Best Local Similarity 77.8%; Pred. No. 94;
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 6 GPAKIOFEG 14
 Db 84 GPAKIOFEG 92

RESULT 15
 US-09-978-697-45
 Sequence 45, Application US/09978697
 Patent No. US20020169284A1
 GENERAL INFORMATION:
 APPLICANT: Ashkenazi, Avi
 APPLICANT: Baker Kevin P.
 APPLICANT: Botstein, David
 APPLICANT: Desnoyers, Luc
 APPLICANT: Eaton, Dan
 APPLICANT: Ferrara, Napoleon
 APPLICANT: Filvaroff, Ellen
 APPLICANT: Fong, Sherman
 APPLICANT: Gao, Wei-Qiang
 APPLICANT: Gerber, Hanspeter
 APPLICANT: Gerlitsen, Mary E.
 APPLICANT: Goddard, Audrey
 APPLICANT: Godowski, Paul J.
 APPLICANT: Grimaldi, J. Christopher
 APPLICANT: Gurney, Austin L.

APPLICANT: Hillan, Kenneth J
APPLICANT: Klijavio, Ivar J.
APPLICANT: Kuo, Sophia S.
APPLICANT: Napier, Mary A.
APPLICANT: Pan, James.
APPLICANT: Paonli, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Shelton, David L.
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: P2630PIC27
CURRENT FILING DATE: 2001-10-16
PRIOR FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: 09/918585
PRIOR FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/064249
PRIOR FILING DATE: 1997-11-03
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066364
PRIOR FILING DATE: 1997-11-21
PRIOR APPLICATION NUMBER: 60/077450
PRIOR FILING DATE: 1998-03-10
PRIOR APPLICATION NUMBER: 60/077632
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077641
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077649
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077791
PRIOR FILING DATE: 1998-03-12
PRIOR APPLICATION NUMBER: 60/078004
PRIOR FILING DATE: 1998-03-13
PRIOR APPLICATION NUMBER: 60/078886
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078936
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078939
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/079294
PRIOR FILING DATE: 1998-03-25
PRIOR APPLICATION NUMBER: 60/079656
PRIOR FILING DATE: 1998-03-26
PRIOR APPLICATION NUMBER: 60/079664
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079689
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079663
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079728
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079786
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079920
PRIOR FILING DATE: 1998-03-30
PRIOR APPLICATION NUMBER: 60/079923
PRIOR FILING DATE: 1998-03-30
PRIOR APPLICATION NUMBER: 60/080105
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080107
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080165
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080194
PRIOR FILING DATE: 1998-03-31

PRIOR APPLICATION NUMBER: 60/080327
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080328
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080333
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080334
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/081070
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081049
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081071
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081195
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081203
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081229
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081955
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081817
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081819
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081952
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081838
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/082568
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082569
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082704
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082804
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082700
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082797
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082796
PRIOR FILING DATE: 1998-04-23
PRIOR APPLICATION NUMBER: 60/083336
PRIOR FILING DATE: 1998-04-27
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/083392
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083495
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083496
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083499
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083545
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083554
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083558
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083559
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083500
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083742
PRIOR FILING DATE: 1998-04-30
PRIOR APPLICATION NUMBER: 60/084366
PRIOR FILING DATE: 1998-05-05
PRIOR APPLICATION NUMBER: 60/084414
PRIOR FILING DATE: 1998-05-06
PRIOR APPLICATION NUMBER: 60/084441

```

; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: 60/084637
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084639
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084640
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084598
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084627
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084643
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/085339
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085338
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085323
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085582
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085700
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085689
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085579
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085580
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085573
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085704
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085697

```

```

Query Match 48.68; Score 34; DB 9; Length 359;
Best Local Similarity 77.88; Pred. No. 94;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 6 GPAKQFEG 14
   |||||
Db 84 GPAKPFEG 92

```

Search completed: January 13, 2003, 10:02:49
Job time : 65.9744 secs

Mon Jan 13 10:35:11 2003

us-09-554-941-11.rapb

Page 1

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 13, 2003, 09:54:16 ; Search time 64.9744 Seconds
(without alignments)
4.180 Million cell updates/sec

Title: US-09-554-941-11
Perfect score: 73
Sequence: 1 GASIRGRTAPDQ 14

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 118974 seqs, 19401057 residues

Total number of hits satisfying chosen parameters: 118974

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published.Applications-AA.*
1: /cgn2_6/ptodata/1/pubppa/US08_NEW_PUB.pep.*
2: /cgn2_6/ptodata/1/pubppa/US06_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubppa/US06_PUBCOMB.pep.*
4: /cgn2_6/ptodata/1/pubppa/US07_NEW_PUB.pep.*
5: /cgn2_6/ptodata/1/pubppa/US07_PUBCOMB.pep.*
6: /cgn2_6/ptodata/1/pubppa/US07_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubppa/US08_PUBCOMB.pep.*
8: /cgn2_6/ptodata/1/pubppa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubppa/US09_NEW_PUB.pep.*
10: /cgn2_6/ptodata/1/pubppa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubppa/US10_NEW_PUB.pep.*
12: /cgn2_6/ptodata/1/pubppa/US10_PUBCOMB.pep.*
13: /cgn2_6/ptodata/1/pubppa/US60_NEW_PUB.pep.*
14: /cgn2_6/ptodata/1/pubppa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	41	56.2	49	10 US-09-864-761-47113	Sequence 47113, A
2	39	53.4	295	10 US-09-893-737-324	Sequence 324, App
3	39	53.4	423	10 US-09-729-674-160	Sequence 160, App
4	39	53.4	428	10 US-09-780-016-18	Sequence 22, Appl
5	39	53.4	453	10 US-09-780-016-18	Sequence 18, Appl
6	39	53.4	484	10 US-09-780-016-26	Sequence 26, Appl
7	39	53.4	507	10 US-09-780-016-2	Sequence 2, Appl
8	39	53.4	509	10 US-09-780-016-24	Sequence 20, Appl
9	39	53.4	532	10 US-09-864-761-39451	Sequence 39451, A
10	39	49.3	70	10 US-09-989-920-184	Sequence 184, App
11	36	49.3	101	9 US-09-794-591-2	Sequence 2, Appl
12	36	49.3	330	10 US-09-995-225-6	Sequence 5433, Ap
13	36	49.3	374	9 US-09-738-626-5433	Sequence 2, Appl
14	36	49.3	407	10 US-09-782-906-2	Sequence 2, Appl
15	35.5	48.6	639	10 US-09-782-906-3	Sequence 4, Appl
16	35.5	48.6	639	10 US-09-782-906-4	Sequence 5, Appl
17	35.5	48.6	639	10 US-09-782-906-5	Sequence 30, Appl
18	35.5	48.6	639	10 US-09-782-906-5	Sequence 30, Appl
19	35	47.9	86	10 US-09-987-108-30	Sequence 30, Appl

20	35	47.9	111	10 US-09-764-869-848	Sequence 848, App
21	35	47.9	460	10 US-09-925-297-584	Sequence 584, App
22	35	47.9	1070	10 US-09-735-367B-6	Sequence 6, Appl
23	35	47.9	2005	10 US-09-735-367B-3	Sequence 3, Appl
24	35	47.9	2063	10 US-09-735-367B-2	Sequence 2, Appl
25	34.5	47.3	651	10 US-09-837-992-3	Sequence 3, Appl
26	34	46.6	147	9 US-09-922-199A-8	Sequence 8, Appl
27	34	46.6	185	10 US-09-922-199A-7	Sequence 7, Appl
28	34	46.6	203	10 US-09-815-242-10084	Sequence 10084, A
29	34	46.6	286	10 US-09-911-826A-4	Sequence 4, Appl
30	34	46.6	309	9 US-09-976-059-34	Sequence 34, Appl
31	34	46.6	1298	9 US-09-825-288A-2	Sequence 2, Appl
32	33.5	45.9	1499	10 US-09-911-826A-2	Sequence 115, App
33	33	45.2	288	10 US-09-888-615-115	Sequence 48476, A
34	33	45.2	34	10 US-09-864-761-48476	Sequence 48910, A
35	33	45.2	34	10 US-09-864-761-48910	Sequence 40180, A
36	33	45.2	68	10 US-09-864-761-40180	Sequence 35131, A
37	33	45.2	127	10 US-09-864-761-35131	Sequence 38443, A
38	33	45.2	141	10 US-09-864-761-38443	Sequence 20, Appl
39	33	45.2	155	9 US-10-014-717-20	Sequence 412, App
40	33	45.2	236	10 US-09-731-872-412	Sequence 1075, Ap
41	33	45.2	244	10 US-09-764-864-1075	Sequence 310, App
42	33	45.2	276	10 US-09-801-368-310	Sequence 434, App
43	33	45.2	280	10 US-09-925-302-454	Sequence 56, Appl
44	33	45.2	283	10 US-09-920-552-56	Sequence 10817, A
45	33	45.2	300	10 US-09-815-242-10817	Sequence 10817, A

ALIGNMENTS

RESULT 1
US-09-864-761-47113
Sequence 47113, Application US/09864761
Patient No. US2002004876341
GENERAL INFORMATION:
APPLICANT: Penn, David R.
APPLICANT: Hank, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME- DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
FILE REFERENCE: Aeonica-X-1
CURRENT APPLICATION NUMBER: US-09/864,761
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30

```

; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 47113
; LENGTH: 49
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC005384.1
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.69
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.77
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.78
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.88
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.79
; OTHER INFORMATION: SWISSPROT HIT: Q04855, EVALU 2.10e+00
US-09-864-761-47113

```

```

Query Match
Best Local Similarity 56.2%; Score 41; DB 10; Length 49;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
QY 4 IEGRYTAPOP 13
:|||||:
DB 20 VEGHRTVPR 29

```

```

RESULT 2
US-09-893-737-324
; Sequence 324, Application US/09893737
; Patent No. US20020110855A1
; GENERAL INFORMATION:
; APPLICANT: Sheppard, Paul O.
; TITLE OF INVENTION: MAMMALIAN SECRETED PROTEINS
; FILE REFERENCE: 00-41
; CURRENT APPLICATION NUMBER: US/09/893,737
; PRIOR FILING DATE: 2001-06-28
; PRIOR APPLICATION NUMBER: US 60/215,446
; PRIOR FILING DATE: 2000-06-30
; NUMBER OF SEQ ID NOS: 329
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 324
; LENGTH: 295
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-893-737-324

```

```

Query Match
Best Local Similarity 53.4%; Score 39; DB 10; Length 295;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 7 RYTAPOP 14
:|||||:
DB 186 KYTAPOPE 193

```

```

RESULT 3
US-09-729-674-160
; Sequence 160, Application US/09729674
; Patent No. US2001003935A1
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Kenneth
; APPLICANT: McCoy, John M.
; APPLICANT: Lavallee, Edward R.
; APPLICANT: Collins-Racie, Lisa A.
; APPLICANT: Evans, Cheryl

```

```

; APPLICANT: Merberg, David
; APPLICANT: Treacy, Maurice
; APPLICANT: Agostino, Michael J.
; APPLICANT: Steininger II, Robert J.
; APPLICANT: Spaulding, Vikki
; APPLICANT: Wong, Gordon G.
; APPLICANT: Clark, Hilary
; APPLICANT: Fectel, Kim
; APPLICANT: Genetics Institute, Inc.
; TITLE OF INVENTION: SECRETED PROTEINS AND POLYPEPTIDES ENCODING THEM
; FILE REFERENCE: 6055-64X
; CURRENT APPLICATION NUMBER: US/09/729,674
; PRIOR FILING DATE: 2000-12-04
; PRIOR APPLICATION NUMBER: 09/539,330
; NUMBER OF SEQ ID NOS: 283
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 160
; LENGTH: 423
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-729-674-160

```

```

Query Match
Best Local Similarity 53.4%; Score 39; DB 10; Length 423;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 7 RYTAPOP 14
:|||||:
DB 88 KYTAPOPE 95

```

```

RESULT 4
US-09-780-016-22
; Sequence 22, Application US/09780016
; Patent No. US20020004591A1
; GENERAL INFORMATION:
; APPLICANT: Donoho, Gregory
; APPLICANT: Scoville, John
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Friedrich, Glenn
; APPLICANT: Abuin, Alejandro
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: No. US20020004591A1 Human Proteases and
; FILE REFERENCE: LEX-0132-USA
; CURRENT APPLICATION NUMBER: US/09/780,016
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: US 60/181,294
; PRIOR FILING DATE: 2000-02-11
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 22
; LENGTH: 428
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-780-016-22

```

```

Query Match
Best Local Similarity 53.4%; Score 39; DB 10; Length 428;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
QY 4 IEGRYTAPOP 14
:|||||:
DB 270 VNGRFTAPOAE 280

```

```

RESULT 5
US-09-780-016-18
; Sequence 18, Application US/09780016
; Patent No. US20020004591A1
; GENERAL INFORMATION:

```

```

; APPLICANT: Donoho, Gregory
; APPLICANT: Scoville, John
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Friedrich, Glenn
; APPLICANT: Abulin, Alejandro
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: Polynucleotides Encoding the Same
; FILE REFERENCE: LEX-0132-USA
; CURRENT APPLICATION NUMBER: US/09/780,016
; PRIOR FILING DATE: 2001-02-09
; PRIOR FILING DATE: 2000-02-11
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: fastseq for Windows Version 4.0
; SEQ ID NO 18
; LENGTH: 453
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-780-016-18

Query Match      53.4%; Score 39; DB 10; Length 453;
Best Local Similarity 54.5%; Pred. No. 28;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 4 IEGRYTAPOQ 14
Db 295 VNGRFTAPOAE 305

RESULT 6
US-09-780-016-26
; Sequence 26, Application US/09780016
; Patent No. US20020004591A1
; GENERAL INFORMATION:
; APPLICANT: Donoho, Gregory
; APPLICANT: Scoville, John
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Friedrich, Glenn
; APPLICANT: Abulin, Alejandro
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: Polynucleotides Encoding the Same
; FILE REFERENCE: LEX-0132-USA
; CURRENT APPLICATION NUMBER: US/09/780,016
; PRIOR FILING DATE: 2001-02-09
; PRIOR FILING DATE: 2000-02-11
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: fastseq for Windows Version 4.0
; SEQ ID NO 26
; LENGTH: 484
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-780-016-26

Query Match      53.4%; Score 39; DB 10; Length 484;
Best Local Similarity 54.5%; Pred. No. 30;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 4 IEGRYTAPOQ 14
Db 326 VNGRFTAPOAE 336

RESULT 7
US-09-780-016-2
; Sequence 2, Application US/09780016
; Patent No. US20020004591A1
; GENERAL INFORMATION:
; APPLICANT: Donoho, Gregory

```

```

; APPLICANT: Scoville, John
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Friedrich, Glenn
; APPLICANT: Abulin, Alejandro
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: Polynucleotides Encoding the Same
; FILE REFERENCE: LEX-0132-USA
; CURRENT APPLICATION NUMBER: US/09/780,016
; PRIOR FILING DATE: 2001-02-09
; PRIOR FILING DATE: 2000-02-11
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: fastseq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 507
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-780-016-2

Query Match      53.4%; Score 39; DB 10; Length 507;
Best Local Similarity 54.5%; Pred. No. 31;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 4 IEGRYTAPOQ 14
Db 349 VNGRFTAPOAE 359

RESULT 8
US-09-780-016-24
; Sequence 24, Application US/09780016
; Patent No. US20020004591A1
; GENERAL INFORMATION:
; APPLICANT: Donoho, Gregory
; APPLICANT: Scoville, John
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Friedrich, Glenn
; APPLICANT: Abulin, Alejandro
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: Polynucleotides Encoding the Same
; FILE REFERENCE: LEX-0132-USA
; CURRENT APPLICATION NUMBER: US/09/780,016
; PRIOR FILING DATE: 2001-02-09
; PRIOR FILING DATE: 2000-02-11
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: fastseq for Windows Version 4.0
; SEQ ID NO 24
; LENGTH: 509
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-780-016-24

Query Match      53.4%; Score 39; DB 10; Length 509;
Best Local Similarity 54.5%; Pred. No. 31;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 4 IEGRYTAPOQ 14
Db 351 VNGRFTAPOAE 361

RESULT 9
US-09-780-016-20
; Sequence 20, Application US/09780016
; Patent No. US20020004591A1
; GENERAL INFORMATION:
; APPLICANT: Donoho, Gregory
; APPLICANT: Scoville, John

```

```
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Friedrich, Glenn
; APPLICANT: Abulin, Alejandro
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: No. US20020004591A1 Human Proteases and
; FILE REFERENCE: LEX-0133-USA
; CURRENT APPLICATION NUMBER: US 09/780,016
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: US 60/181,294
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 20
; LENGTH: 532
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-780-016-20
```

```
Query Match
Best Local Similarity 53.4%; Score 39; DB 10; Length 532;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
```

```
QY 4 IEGRYTAPQPO 14
Db 374 VNGRTAPQAE 384
```

```
RESULT 10
US-09-864-761-39451
; Sequence 39451, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aecmca-x-1
; CURRENT APPLICATION NUMBER: US 09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263,6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
```

```
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 39451
; LENGTH: 70
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC006075.1
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.1
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.9
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.1
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.3
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.6
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.9
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.9
; OTHER INFORMATION: SWISSPROT HIT: Q09457, EVALUATE 5.00e-03
; OTHER INFORMATION: EST_HUMAN HIT: AUI18435.1, EVALUATE 8.00e-37
US-09-864-761-39451
```

```
Query Match
Best Local Similarity 49.3%; Score 36; DB 10; Length 70;
Matches 7; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
```

```
QY 1 QASIEGRYTPQPO 13
Db 14 QNGCPAETRYTPHP 26
```

```
RESULT 11
US-09-989-920-184
; Sequence 184, Application US/09989920
; Patent No. US20020172957A1
; GENERAL INFORMATION:
; APPLICANT: Macina, Roberto
; APPLICANT: Recipon, Hervé
; APPLICANT: Chen, Sel-Yu
; APPLICANT: Sun, Yongming
; APPLICANT: Liu, Chenghua
; TITLE OF INVENTION: Compositions and Methods Relating to Lung Specific Genes and P
; FILE REFERENCE: DEX-0291
; CURRENT APPLICATION NUMBER: US 09/989,920
; CURRENT FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: 60/252,500
; PRIOR FILING DATE: 2000-11-22
; NUMBER OF SEQ ID NOS: 284
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 184
; LENGTH: 101
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-989-920-184
```

```
Query Match
Best Local Similarity 49.3%; Score 36; DB 9; Length 101;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 6 GRYTAPQ 13
Db 4 GRYQAPRP 11
```

```
RESULT 12
US-09-794-591-2
; Sequence 2, Application US/09794591
; Patent No. US20010018198A1
; GENERAL INFORMATION:
```

APPLICANT: Pulst, Stefan M.
APPLICANT: Shidota, Hiroki
TITLE OF INVENTION: Nucleic Acids Encoding Ataxin-2 Binding Proteins,
TITLE OF INVENTION: Products Related Thereto and Methods of Using Same
FILE REFERENCE: CE 3093
CURRENT APPLICATION NUMBER: US/09/794,591
PRIOR FILING DATE: 2001-02-26
PRIOR APPLICATION NUMBER: 09/145,391
NUMBER OF SEQ ID NOS: 2
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 2
LENGTH: 330
TYPE: PRT
ORGANISM: Homo sapiens
US-09-794-591-2

Query Match
Best Local Similarity 49.3%; Score 36; DB 10; Length 330;
Matches 7; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 QASIEGRYTAPOP 13
DB 34 QNGIPAEYTAAPH 46

RESULT 13
US-09-995-225-6
Sequence 6, Application US/0995225
Publication No. US20020193584A1
GENERAL INFORMATION:
APPLICANT: Chen, Ruoping
APPLICANT: Chu, Zhi Liang
APPLICANT: Dang, Huang T.
APPLICANT: Lowitz, Kevin P.
APPLICANT: Pride, Cameron
TITLE OF INVENTION: Endogenous And No. US20020193584A1-Endogenous Versions of Human G
TITLE OF INVENTION: Receptors
FILE REFERENCE: AREN-0308
CURRENT APPLICATION NUMBER: US/09/995,225
CURRENT FILING DATE: 2001-11-26
PRIOR APPLICATION NUMBER: 09/170,496
PRIOR FILING DATE: 1998-10-13
PRIOR APPLICATION NUMBER: PCT/US99/23938
PRIOR FILING DATE: 1998-10-13
PRIOR APPLICATION NUMBER: 60/253,404
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/255,366
PRIOR FILING DATE: 2000-12-12
PRIOR APPLICATION NUMBER: 60/270,286
PRIOR FILING DATE: 2001-02-20
PRIOR APPLICATION NUMBER: 60/282,365
PRIOR FILING DATE: 2001-04-06
PRIOR APPLICATION NUMBER: 60/270,266
PRIOR FILING DATE: 2001-02-20
PRIOR APPLICATION NUMBER: 60/282,032
PRIOR FILING DATE: 2001-04-06
PRIOR APPLICATION NUMBER: 60/282,358
PRIOR FILING DATE: 2001-04-06
PRIOR APPLICATION NUMBER: 60/282,356
PRIOR FILING DATE: 2001-04-06
PRIOR APPLICATION NUMBER: 60/290,917
PRIOR FILING DATE: 2001-05-14
PRIOR APPLICATION NUMBER: 60/309,208
PRIOR FILING DATE: 2001-07-31
NUMBER OF SEQ ID NOS: 67
SOFTWARE: PatentIn version 3.1
SEQ ID NO 6
LENGTH: 374
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: No. US20020193584A1 Sequence

US-09-995-225-6

Query Match
Best Local Similarity 49.3%; Score 36; DB 9; Length 374;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 4 IEGRYTAPOP 14
DB 26 LEGRETAGOPR 36

RESULT 14
US-09-738-626-5433
Sequence 5433, Application US/09738626
Publication No. US20020197605A1
GENERAL INFORMATION:
APPLICANT: NAKAGAWA, SATOSHI
APPLICANT: MIZOGUCHI, HIROSHI
APPLICANT: ANDO, SEIKO
APPLICANT: HAYASHI, MIKIRO
APPLICANT: OCHIAI, KEIKO
APPLICANT: YOKOI, HARUHIKO
APPLICANT: TATEISHI, NAOKO
APPLICANT: SENOH, AKIHIRO
APPLICANT: IKEDA, MASATO
APPLICANT: OZAKI, AKIO
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-125
CURRENT APPLICATION NUMBER: US/09/738,626
CURRENT FILING DATE: 2000-12-18
PRIOR APPLICATION NUMBER: JP 99/377484
PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: JP 00/159162
PRIOR FILING DATE: 2000-04-07
PRIOR APPLICATION NUMBER: JP 00/280988
PRIOR FILING DATE: 2000-08-03
NUMBER OF SEQ ID NOS: 7059
SOFTWARE: PatentIn ver. 3.0
SEQ ID NO 5433
LENGTH: 407
TYPE: PRT
ORGANISM: Corynebacterium glutamicum
US-09-738-626-5433

Query Match
Best Local Similarity 49.3%; Score 36; DB 9; Length 407;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 2 ASIEGRYTAPOP 14
DB 248 AAAPQOYAPOPQ 260

RESULT 15
US-09-782-906-2
Sequence 2, Application US/09782906
Patent No. US20010051369A1
GENERAL INFORMATION:
APPLICANT: Delagrave, Simon
APPLICANT: Rittenhouse Pruss, Jennifer L.
APPLICANT: Murphy, Dennis J.
APPLICANT: Maffia III, Anthony M.
APPLICANT: Bylina, Edward J.
APPLICANT: Coleman, William J.
TITLE OF INVENTION: Variant Galactose Oxidase, Nucleic Acid Encoding Same, And Met
FILE REFERENCE: HER-0040
CURRENT APPLICATION NUMBER: US/09/782,906
CURRENT FILING DATE: 2001-02-14
PRIOR APPLICATION NUMBER: 60/185,001
PRIOR FILING DATE: 2000-02-25
NUMBER OF SEQ ID NOS: 8
SOFTWARE: PatentIn version 3.0

; SEQ ID NO 2
; LENGTH: 639
; TYPE: PRT
; ORGANISM: Dactylium dendroides
US-09-782-906-2

Query Match 48.6%; Score 35.5; DB 10; Length 639;
Best Local Similarity 69.2%; Pred. No. 1.6e+02;
Matches 9; Conservative 0; Mismatches 1; Indels 3; Gaps 1;

OY 1 QASIEGRYTAPOP 13
Db 147 QAS---SYTAPOP 156

Search completed: January 13, 2003, 10:02:50
Job time : 65.9744 secs

Mon Jan 13 10:35:11 2003

us-09-554-941-11.rpr

Page 1

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 13, 2003, 09:53:20 ; Search time 12.9231 seconds
(without alignments)
104.146 Million cell updates/sec

Title: US-09-554-941-11
Perfect score: 73
Sequence: 1 QASIEGRYTAPOPQ 14

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	45	61.6	326	A59232	ABA-responsive pro
2	43	58.9	295	JC6530	laminin receptor p
3	42	57.5	424	S76359	hypothetical prote
4	41	56.2	224	B75388	NADH dehydrogenase
5	40	54.8	1263	AH2011	heterocyst glycoli
6	39	53.4	362	S78515	single-stranded nu
7	39	53.4	365	S42471	hnRNP protein E2 -
8	39	53.4	506	T35923	probable tRNA synt
9	39	53.4	508	A33378	fasciclin III prec
10	39	53.4	649	G66434	protein F17F8.23 l
11	39	53.4	1571	T00062	hypothetical prote
12	38	52.1	75	C81322	probable periplasm
13	38	52.1	271	T24965	hypothetical prote
14	38	52.1	293	AC3464	virulence-associat
15	38	52.1	628	T02602	vacuolar sorting r
16	38	52.1	628	T02602	probable vacuolar
17	38	52.1	662	T02604	transposase - phag
18	38	52.1	3942	T42730	Bassoon protein -
19	38	52.1	786	D75630	glycerophosphoryl
20	37.5	51.4	138	SWMOD	diuretic hormone p
21	37	50.7	156	T29730	hypothetical prote
22	37	50.7	286	AF0610	probable p17n-rel
23	37	50.7	288	AE0056	probable right ori
24	37	50.7	296	C49539	xyloglucan endo-1,
25	37	50.7	396	D83520	hypothetical prote
26	37	50.7	401	F69634	multidrug-efflux t
27	37	50.7	478	T51274	hypothetical prote
28	37	50.7	608	T02299	hypothetical prote
29	37	50.7	623	T47542	Spot 3 protein and

30	37	50.7	623	2	T06794	vacuolar sorting r
31	37	50.7	624	2	T00044	vacuolar sorting r
32	37	50.7	625	2	F84706	probable vacuolar
33	37	50.7	626	2	T04895	vacuolar sorting r
34	37	50.7	768	2	I53821	P-selectin - rat
35	37	50.7	768	2	A42755	P-selectin precurs
36	37	50.7	798	2	S20681	homeotic protein p
37	37	50.7	961	2	JC2200	alpha-mannosidase
38	37	50.7	1116	2	D97695	COBN protein homol
39	37	50.7	1116	2	AC2921	copalamin biosynth
40	37	50.7	1124	2	T30340	dsRNA adenosine de
41	37	50.7	2219	2	T27684	hypothetical prote
42	37	50.7	4385	2	T29042	hypothetical prote
43	36.5	50.0	337	2	T32142	hypothetical prote
44	36	49.3	22	2	I51566	protein-tyrosine k
45	36	49.3	176	2	AE0526	2'-5' RNA ligase (

ALIGNMENTS

RESULT 1
A59232
ABA-responsive protein - barley
C:Species: Hordeum vulgare (barley)
C>Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text-change 02-Jun-2000
C:Accession: A59232
R:Jiliu, J.H.; Luo, M.; Mohapatra, S.S.
submitted to GenBank, September 1997
A:Reference number: A59232
A:Accession: A59232
A>Status: not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-326 <LRU>
A:Cross-references: GB:AF026538; NID:94103634; PIDN:AA009343.1; PID:94103635
A:Experimental source: dev stage embryo
C:Superfamily: barley ABA-responsive protein

Query Match
Best Local Similarity 70.0%; Score 45; DB 2; Length 326;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 4 IEGRYTAPOP 13
DB 76 IEGRYTAPOP 85

RESULT 2
JC6530
laminin receptor processed pseudogene LAMRL5 - human
C:Species: Homo sapiens (man)
C>Date: 28-Aug-1998 #sequence_revision 28-Aug-1998 #text-change 28-Aug-1998
C:Accession: JC6530
R:Richardson, M.P.; Braybrook, C.; Tham, M.; Moore, G.E.; Stanier, P.
Gene 206, 145-150, 1998
A>Title: Molecular cloning and characterization of a highly conserved human 67-kDa la
A:Reference number: JC6530; MUID:98121324; PMID:9461426
A:Accession: JC6530
A:Status: conceptual translation of pseudogene
A:Molecule type: DNA
A:Residues: 1-295 <RIC>
A:Experimental source: brain
C:Comment: No evidence could be found that this intronless gene sequence is expressed

A:Gene: LAMRL5
A:Map position: Xq21.3
A:Introns: #status absent
C:Keywords: brain; glycoprotein; laminin binding; pseudogene; receptor

Query Match
Best Local Similarity 63.6%; Score 43; DB 4; Length 295;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 4 IEGRYTAPOP 14
 DB 227 LEGOWTAPAPQ 237

RESULT 3

S76359
 hypothetical protein - *Synechocystis* sp. (strain PCC 6803)

C:Species: *Synechocystis* sp.
 A:Variate: PCC 6803
 C:Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 20-Jun-2000
 C:Accession: S76359
 O:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.; DNA Res. 3, 109-136, 1996
 A:Title: Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis*
 A:Reference number: S74322; MUID:97061201; PMID:8905231
 A:Accession: S76359
 A:Status: Preliminary
 A:Molecule type: DNA
 A:Residues: 1-424 <KAN>
 A:Cross-references: EMBL:D64000; GB:AB001339; NID:91001484; PIDN:BA110211.1; PID:9100158
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996
 C:Superfamily: site-specific methyltransferase (cytosine-specific) EcoRII

Query Match
 Best Local Similarity 57.5%; Score 42; DB 2; Length 424;
 Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 2 ASIEGRTAPOP 13
 DB 319 ASDGATAPAPR 330

RESULT 4

B75388

NADH dehydrogenase I, C subunit - *Deinococcus radiodurans* (strain R1)

C:Species: *Deinococcus radiodurans*
 C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 31-Mar-2000
 C:Accession: B75388
 R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.; S.; Smith, H.O.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Uitterback, T.; Zalewski, C.; M. Science 286, 1571-1577, 1999
 A:Title: Genome sequence of the radioresistant bacterium *Deinococcus radiodurans* R1.
 A:Reference number: A75250; MUID:20036896; PMID:10567266
 A:Accession: B75388
 A:Status: Preliminary
 A:Molecule type: DNA
 A:Residues: 1-224 <WHI>
 A:Cross-references: GB:AE001994; GB:AE000513; NID:96459259; PIDN:AAF11071.1; PID:9645926
 C:Genetics:
 A:Gene: DR1504
 A:Map position: 1

Query Match
 Best Local Similarity 56.2%; Score 41; DB 2; Length 224;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 7 RYTAPOP 14
 DB 80 RYTAPOPQ 87

RESULT 5

AH2011

heterocyst glycolipid synthase [Imported] - *Nostoc* sp. (strain PCC 7120)

C:Species: *Nostoc* sp.
 A:Note: *Nostoc* sp. strain PCC 7120 is a synonym of *Anabaena* sp. strain PCC 7120
 C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 30-Jun-2002
 C:Accession: AH2011

R.Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kunitz, T.; Sasamoto, S.; Watanabe, A.; Irigu
 Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata
 DNA Res. 8, 205-213, 2001
 A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium
 A:Reference number: AB1807; MUID:21595285; PMID:11759840
 A:Accession: AH2011
 A:Status: Preliminary
 A:Molecule type: DNA
 A:Residues: 1-1263 <KAN>
 A:Cross-references: GB:BA000019; PIDN:BA878012.1; PID:911135466; GSPDB:GN00179
 A:Experimental source: strain PCC 7120
 C:Genetics:
 A:Gene: hglE

Query Match
 Best Local Similarity 54.8%; Score 40; DB 2; Length 1263;
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 4 IEGRYTAPOP 14
 DB 1154 LEPRYSAPAPQ 1164

RESULT 6

S78515

single-stranded nucleic acid-binding protein CBP - mouse

N:Alternate names: hnRNP protein X
 C:Species: *Mus musculus* (house mouse)
 C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 05-Nov-1999
 C:Accession: S78515; S45261; S78514; S41652; I48281
 R:Hahn, K.B.; Kim, G.; Turch, C.; Smale, S.T.
 A:Reference number: 148281; MUID:94268912; PMID:8208614
 A:Accession: S78515
 A:Status: nucleic acid sequence not shown
 A:Molecule type: mRNA
 A:Residues: 1-362 <HAN>
 A:Cross-references: EMBL:L19661; NID:9436893; PIDN:AAA03705.1; PID:9436894
 R:Goller, M.; Funke, B.; Gene-Becker, C.; Kroege, B.; Lottspeich, F.; Horak, I.
 Nucleic Acids Res. 22, 1885-1889, 1994
 A:Title: Murine protein which binds preferentially to oligo-C-rich single-stranded nu
 A:Reference number: 148281; MUID:94268912; PMID:8208614
 A:Accession: S45261
 A:Status: nucleic acid sequence not shown
 A:Molecule type: mRNA
 A:Residues: 1-193, 225-300, 302-362 <GOL>
 A:Cross-references: EMBL:X75947; NID:9495127; PIDN:CAA53546.1; PID:9495128
 A:Accession: S78514
 A:Molecule type: protein
 A:Residues: 12-39; 172-193; 225-229; 306-322 <GOW>
 R:Hahn, K.; Kim, G.; Turch, C.W.; Smale, S.T.
 Nucleic Acids Res. 21, 3894, 1993
 A:Title: Isolation of a murine gene encoding a nucleic acid-binding protein with homo
 A:Reference number: S41652; MUID:93376518; PMID:8367306
 A:Accession: S41652
 A:Molecule type: mRNA
 A:Residues: 16-51; 100-135; 287-322 <HAN>
 A:Genetics:
 A:Gene: CBP
 C:Keywords: alternative splicing; RNA binding; single-stranded DNA binding
 F:1-362/Product: single-stranded nucleic acid-binding protein CBP, long form #status
 F:1-193, 225-362/Product: single-stranded nucleic acid-binding protein CBP, short form

Query Match
 Best Local Similarity 53.4%; Score 39; DB 2; Length 362;
 Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 3 SIEGRTAPOP 13
 DB 227 TIQGYTAPOP 237

RESULT 7

S42471

hnpnp protein f2 - human

N:Alternate names: nucleic acid-binding protein; protein PCBP-2

C:Species: Homo sapiens (man)

C>Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 05-Nov-1999

C:Accession: S65679; S42471

R:Leffers, H.; Dejgaard, K.; Celis, J.E.

Eur. J. Biochem. 230, 447-453, 1995

A:Title: Characterisation of two major cellular poly(rC)-binding human proteins, each cc

A:Reference number: S65678; MUID:93331278; PMID:7607214

A:Accession: S65679

A:Status: nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 1-365 <LEP>

A:Cross-references: EMBL:X78136; NID:9460772; PIDN:CAA55015.1; PID:9460773

A:Experimental source: AMA cells (transformed human amnion cells)

A:Note: submitted to the EMBL Data Library, March 1994

C:Keywords: RNA binding

Query Match 53.4%; Score 39; DB 2; Length 365;

Best Local Similarity 54.5%; Pred. No. 32;

Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 3 SIEGRYTAPOP 13

Db 231 TIGGYAIRPOP 241

RESULT 8

probable tRNA synthetase - Streptomyces coelicolor

C:Species: Streptomyces coelicolor

C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 17-Mar-2000

C:Accession: T35923

R:Seeger, K.J.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.

submitted to the EMBL Data Library, January 1999

A:Reference number: Z11551

A:Accession: T35923

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-506 <SEP>

A:Cross-references: EMBL:AL035206; PIDN:CAA22745.1; GSPDB:GN00070; SCOEDB:SC9B5.03

A:Experimental source: strain A3(2)

C:Genetics:

A:Gene: SCOEDB:SC9B5.03

C:Superfamily: methionine-tRNA ligase

Query Match 53.4%; Score 39; DB 2; Length 506;

Best Local Similarity 58.3%; Pred. No. 45;

Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 ASTEGRYTAPOP 13

Db 481 AAVEARTTAAP 492

RESULT 9

faciclin III precursor - fruit fly (Drosophila melanogaster)

C:Species: Drosophila melanogaster

C>Date: 21-Feb-1990 #sequence_revision 21-Feb-1990 #text_change 24-Sep-1998

C:Accession: A33378

R:Snow, P.M.; Bieber, A.J.; Goodman, C.S.

Cell 59, 313-323, 1989

A:Title: Faciclin III: a novel homophilic adhesion molecule in Drosophila.

A:Reference number: A33378; MUID:90030406; PMID:2509076

A:Accession: A33378

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-508 <SNO>

A:Cross-references: GB:M27813; NID:9157423; PID:9157424

C:Genetics:

A:Gene: FlyBase:FBgn0000636

A:Cross-references: FlyBase:FBgn0000636

C:Keywords: phosphoprotein; transmembrane protein

Query Match 53.4%; Score 39; DB 2; Length 508;

Best Local Similarity 50.0%; Pred. No. 45;

Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 QASIEGRYTAPOP 14

Db 279 QGRIDGRISATPEPQ 292

RESULT 10

protein F17F8.23 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001

C:Accession: G86434

R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alon

Chiu, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar,

ansen, N.F.; Hughes, B.; Huizar, L.

Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marzla

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallo

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A:Reference number: A86141; MUID:21016719; PMID:11130712

A:Accession: G86434

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-649 <STOP>

A:Cross-references: GB:AE005172; NID:99755389; PIDN:AAF98196.1; GSPDB:GN00141

C:Genetics:

A:Gene: F17F8.23

A:Map position: 1

Query Match 53.4%; Score 39; DB 2; Length 649;

Best Local Similarity 60.0%; Pred. No. 59;

Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 5 EGRYTAPOPQ 14

Db 255 QGRYCAPDPE 264

RESULT 11

hypothetical protein KIAA0434 - human (fragment)

C:Species: Homo sapiens (man)

C>Date: 22-Jan-1999 #sequence_revision 22-Jan-1999 #text_change 23-Mar-2001

C:Accession: T00062

R:Rishikawa, K.; Nagase, T.; Nakajima, D.; Seki, N.; Ohira, M.; Miyajima, N.; Tanaka,

submitted to the EMBL Data Library, October 1997

A:Description: Prediction of the coding sequences of unidentified human genes. VIII.

A:Reference number: Z14082

A:Accession: T00062

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-1571 <ISH>

A:Cross-references: EMBL:AB007894; NID:92662148; PIDN:BA23707.1; PID:92662149

A:Experimental source: brain; clone HH2165

C:Genetics:

A:Note: KIAA0434

Query Match 53.4%; Score 39; DB 2; Length 1571;

Best Local Similarity 50.0%; Pred. No. 1.5e+02;

Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 QASIEGRYTAPOPQ 14

Db 1389 QPOLGROAAPEPQ 1402

RESULT 12
 C81322
 probable periplasmic protein Cj1169c [imported] - Campylobacter jejuni (strain NCTC 1116)
 C:Species: Campylobacter jejuni
 C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 03-Jun-2002
 C:Accession: C81322
 R:Parkhill, J.; Wren, B.W.; Mungall, K.; Kelley, J.M.; Churcher, C.; Basham, D.; Chilling
 C.W.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVliet, A.; Whitehead, S.; Barrell
 Nature 403, 665-668, 2000
 A:Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hyf
 A:Reference number: AB1250; MUID:20150912; PMID:10686204
 A:Accession: C81322
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-75 <PAR>
 A:Cross-references: GB:AL139077; GB:AL111168; NID:g6568444; PIDN:CAW7423.1; PID:g656860
 A:Experimental source: serotype O2, strain NCTC 11168
 C:Genetics:
 A:Gene: Cj1169c

Query Match 52.1%; Score 38; DB 2; Length 75;
 Best Local Similarity 46.2%; Pred. No. 8, 6;
 Matches 6; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

OY 2 ASIEGRTAPOPQ 14
 ||: ||: ||: ||:
 Db 19 ASIMAKYEAPPE 31

RESULT 13
 T24965
 hypothetical protein T18D3.6 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000
 C:Accession: T24965
 R:McMurray, A.
 Submitted to the EMBL Data Library, November 1995
 A:Reference number: Z19962
 A:Accession: T24965
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-271 <MIL>
 A:Cross-references: EMBL:Z68119; PIDN:CAA92195.1; GSPDB:GN00028; CESP:T18D3.6
 A:Experimental source: clone T18D3
 C:Genetics:
 A:Gene: CESP:T18D3.6
 A:Map position: X
 A:Introns: 247/3

Query Match 52.1%; Score 38; DB 2; Length 271;
 Best Local Similarity 70.0%; Pred. No. 35;
 Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 5 EGRYTAPOPO 14
 |||| ||:
 Db 33 EGRYLPPLPQ 42

RESULT 14
 AC3464
 virulence-associated protein E [imported] - Brucella melitensis (strain 16M)
 C:Species: Brucella melitensis
 C:Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 15-Feb-2002
 C:Accession: AC3464
 R:DeIvecchio, V.G.; Kapatal, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova,
 .; Mazur, M.; Goldsman, E.; Selkov, E.; Elizer, P.H.; Hagius, S.; O'Callaghan, D.; Letess
 Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
 A:Title: The genome sequence of the facultative intracellular pathogen Brucella melitens
 A:Reference number: AD3252; PMID:11756688
 A:Accession: AC3464
 A:Status: preliminary
 A:Molecule type: DNA

A:Residues: 1-293 <KUR>
 A:Cross-references: GB:AE008917; PIDN:AA152878.1; PID:g17983722; GSPDB:GN00190
 A:Experimental source: strain 16M
 C:Genetics:
 A:Gene: BME11697
 A:Map position: I
 C:Superfamily: Sphingomonas aromaticivorans hypothetical protein 229

Query Match 52.1%; Score 38; DB 2; Length 293;
 Best Local Similarity 60.0%; Pred. No. 38;
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 5 EGRYTAPOPO 14
 |||| ||:
 Db 70 EGRYTPSPPE 79

RESULT 15
 T02602
 vacuolar sorting receptor protein homolog At2g14740 - Arabidopsis thaliana
 M:Alternate names: vacuolar sorting receptor protein homolog F26C24.12
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 16-Feb-2001
 C:Accession: T02602; H84520
 R:Rounsley, S.D.; Kaul, S.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Syke
 submitted to the EMBL Data Library, June 1998
 A:Description: Arabidopsis thaliana chromosome II BAC F26C24 genomic sequence.
 A:Reference number: Z14680
 A:Accession: T02602
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-628 <ROU>
 A:Cross-references: EMBL:AC004705; NID:g3252804; PID:g3252813
 R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.L.; Town, C.D.; Fujii, C.Y
 M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon,
 euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter
 Nature 402, 761-768, 1999
 A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
 A:Reference number: AB4420; MUID:20083487; PMID:10617197
 A:Accession: H84520
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-628 <STO>
 A:Cross-references: GB:AE002093; NID:g3252813; PIDN:AMC24183.1; GSPDB:GN00139
 C:Genetics:
 A:Gene: F26C24.12; AT2g14740
 A:Map position: 2
 A:Introns: 102/1; 333/1; 362/3; 386/1; 414/1; 439/3; 467/1; 494/3; 517/1; 560/1; 596/

Query Match 52.1%; Score 38; DB 2; Length 628;
 Best Local Similarity 60.0%; Pred. No. 86;
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 5 EGRYTAPOPO 14
 :||| ||:
 Db 258 KGRYCAPDPE 267

Search completed: January 13, 2003, 09:57:11
 Job time : 14.9231 secs

Gencore version 5.1.3
© 2003 Compugen Ltd.

Copyright (c) 1993 - 2000

...in search, using sw model

OM protein - protein search, using sw module - Search time 53.1282 seconds
 January 13, 2003, 09:53:20 (without alignments)
 Run on: 54.256 Million cell updates/sec

US-09-554-941-11

File 73
perfect score: 1
QASIEGRTAPQPQ 14

Sequence: 1 QASLEON

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues
Chosen parameters: 671580

Minimum DB seq length:	0
Maximum DB seq length:	200000000000

Minimum DB seq	length:	2000000000
Maximum DB seq		

Minimum	Match 08
Maximum	Match 100%
post-processing:	
Minimum	Match 08
Maximum	Match 100%
45 summaries	

```
Database : SPTREMBL_21:*
1: sp_archaea:*
```

```

2: sp_bacteria.*
3: sp_fungi.*
4: sp_human.*
5: sp_invertebrate.*
6: sp_mammal.*
7: sp_mhc.*
8: sp_organelle.*
9: sp_phage.*
10: sp_plant.*
11: sp_rodent.*
12: sp_virus.*
13: sp_vertebrate.*
14: sp_unclassified.*
15: sp_virus.*
16: sp_bacteriap.*
17: sp_archaeap.*

```

17: sp_arcneyf.

NO. is the number of results predicted by the normal distribution. The score of the result being printed.

pred. No. is the number of observations greater than or equal to the score of the total score distribution.

score greater than or equal to the total score distribution is derived by analysis of the total score distribution.

score 5.5 and is derived by analysis of variance. SUMMARY

SUMMARIES

[illegible]

Query	Description
8	

Query	DB	ID
length	borderw	000000

Score	Match	Length	DB	ID
0.99970	Q9ZTW0	hordeum		
0.99970	Q9ZTW0	droseroph		

Score	Success	Q9ZTW0
326	10	Q9ZTW0
616		Q9VZr0 drosoph
		Q9VZr0 drosoph

1	45	61.6	326	10	Q9Z1W0	Q9Z1W0	Q9SV2	grosoph
		60.0	283	5	Q9VZR0		Q9SV2	grosoph

Strain	Genotype	Genotype	Genotype
1	45	01.0	Q95SV2
2	43	58.9	Q9Kzd6 strept
3	43	58.9	Q9Kzd6 strept
4	43	58.9	Q9Kzd6 strept
5	43	58.9	Q9Kzd6 strept
6	43	58.9	Q9Kzd6 strept
7	43	58.9	Q9Kzd6 strept
8	43	58.9	Q9Kzd6 strept
9	43	58.9	Q9Kzd6 strept
10	43	58.9	Q9Kzd6 strept
11	43	58.9	Q9Kzd6 strept
12	43	58.9	Q9Kzd6 strept
13	43	58.9	Q9Kzd6 strept
14	43	58.9	Q9Kzd6 strept
15	43	58.9	Q9Kzd6 strept
16	43	58.9	Q9Kzd6 strept
17	43	58.9	Q9Kzd6 strept
18	43	58.9	Q9Kzd6 strept
19	43	58.9	Q9Kzd6 strept
20	43	58.9	Q9Kzd6 strept
21	43	58.9	Q9Kzd6 strept
22	43	58.9	Q9Kzd6 strept
23	43	58.9	Q9Kzd6 strept
24	43	58.9	Q9Kzd6 strept
25	43	58.9	Q9Kzd6 strept
26	43	58.9	Q9Kzd6 strept
27	43	58.9	Q9Kzd6 strept
28	43	58.9	Q9Kzd6 strept
29	43	58.9	Q9Kzd6 strept
30	43	58.9	Q9Kzd6 strept
31	43	58.9	Q9Kzd6 strept
32	43	58.9	Q9Kzd6 strept
33	43	58.9	Q9Kzd6 strept
34	43	58.9	Q9Kzd6 strept
35	43	58.9	Q9Kzd6 strept
36	43	58.9	Q9Kzd6 strept
37	43	58.9	Q9Kzd6 strept
38	43	58.9	Q9Kzd6 strept
39	43	58.9	Q9Kzd6 strept
40	43	58.9	Q9Kzd6 strept
41	43	58.9	Q9Kzd6 strept
42	43	58.9	Q9Kzd6 strept
43	43	58.9	Q9Kzd6 strept
44	43	58.9	Q9Kzd6 strept
45	43	58.9	Q9Kzd6 strept
46	43	58.9	Q9Kzd6 strept
47	43	58.9	Q9Kzd6 strept
48	43	58.9	Q9Kzd6 strept
49	43	58.9	Q9Kzd6 strept
50	43	58.9	Q9Kzd6 strept
51	43	58.9	Q9Kzd6 strept
52	43	58.9	Q9Kzd6 strept
53	43	58.9	Q9Kzd6 strept
54	43	58.9	Q9Kzd6 strept
55	43	58.9	Q9Kzd6 strept
56	43	58.9	Q9Kzd6 strept
57	43	58.9	Q9Kzd6 strept
58	43	58.9	Q9Kzd6 strept
59	43	58.9	Q9Kzd6 strept
60	43	58.9	Q9Kzd6 strept
61	43	58.9	Q9Kzd6 strept
62	43	58.9	Q9Kzd6 strept
63	43	58.9	Q9Kzd6 strept
64	43	58.9	Q9Kzd6 strept
65	43	58.9	Q9Kzd6 strept
66	43	58.9	Q9Kzd6 strept
67	43	58.9	Q9Kzd6 strept
68	43	58.9	Q9Kzd6 strept
69	43	58.9	Q9Kzd6 strept
70	43	58.9	Q9Kzd6 strept
71	43	58.9	Q9Kzd6 strept
72	43	58.9	Q9Kzd6 strept
73	43	58.9	Q9Kzd6 strept
74	43	58.9	Q9Kzd6 strept
75	43	58.9	Q9Kzd6 strept
76	43	58.9	Q9Kzd6 strept
77	43	58.9	Q9Kzd6 strept
78	43	58.9	Q9Kzd6 strept
79	43	58.9	Q9Kzd6 strept
80	43	58.9	Q9Kzd6 strept
81	43	58.9	Q9Kzd6 strept
82	43	58.9	Q9Kzd6 strept
83	43	58.9	Q9Kzd6 strept
84	43	58.9	Q9Kzd6 strept
85	43	58.9	Q9Kzd6 strept
86	43	58.9	Q9Kzd6 strept
87	43	58.9	Q9Kzd6 strept
88	43	58.9	Q9Kzd6 strept

Order	Family	Species	Count	Percentage
1	Q95SV2	Q95SV2	5	323
2	Q9KZD6	Q9KZD6	16	525
3	Q9KZD6	Q9KZD6	16	525
4	Q9KZD6	Q9KZD6	16	525
5	Q9KZD6	Q9KZD6	16	525
6	Q9KZD6	Q9KZD6	16	525
7	Q9KZD6	Q9KZD6	16	525
8	Q9KZD6	Q9KZD6	16	525
9	Q9KZD6	Q9KZD6	16	525
10	Q9KZD6	Q9KZD6	16	525
11	Q9KZD6	Q9KZD6	16	525
12	Q9KZD6	Q9KZD6	16	525
13	Q9KZD6	Q9KZD6	16	525
14	Q9KZD6	Q9KZD6	16	525
15	Q9KZD6	Q9KZD6	16	525
16	Q9KZD6	Q9KZD6	16	525
17	Q9KZD6	Q9KZD6	16	525
18	Q9KZD6	Q9KZD6	16	525
19	Q9KZD6	Q9KZD6	16	525
20	Q9KZD6	Q9KZD6	16	525
21	Q9KZD6	Q9KZD6	16	525
22	Q9KZD6	Q9KZD6	16	525
23	Q9KZD6	Q9KZD6	16	525
24	Q9KZD6	Q9KZD6	16	525
25	Q9KZD6	Q9KZD6	16	525
26	Q9KZD6	Q9KZD6	16	525
27	Q9KZD6	Q9KZD6	16	525
28	Q9KZD6	Q9KZD6	16	525
29	Q9KZD6	Q9KZD6	16	525
30	Q9KZD6	Q9KZD6	16	525
31	Q9KZD6	Q9KZD6	16	525
32	Q9KZD6	Q9KZD6	16	525
33	Q9KZD6	Q9KZD6	16	525
34	Q9KZD6	Q9KZD6	16	525
35	Q9KZD6	Q9KZD6	16	525
36	Q9KZD6	Q9KZD6	16	525
37	Q9KZD6	Q9KZD6	16	525
38	Q9KZD6	Q9KZD6	16	525
39	Q9KZD6	Q9KZD6	16	525
40	Q9KZD6	Q9KZD6	16	525
41	Q9KZD6	Q9KZD6	16	525
42	Q9KZD6	Q9KZD6	16	525
43	Q9KZD6	Q9KZD6	16	525
44	Q9KZD6	Q9KZD6	16	525
45	Q9KZD6	Q9KZD6	16	525
46	Q9KZD6	Q9KZD6	16	525
47	Q9KZD6	Q9KZD6	16	525
48	Q9KZD6	Q9KZD6	16	525
49	Q9KZD6	Q9KZD6	16	525
50	Q9KZD6	Q9KZD6	16	525
51	Q9KZD6	Q9KZD6	16	525
52	Q9KZD6	Q9KZD6	16	525
53	Q9KZD6	Q9KZD6	16	525
54	Q9KZD6	Q9KZD6	16	525
55	Q9KZD6	Q9KZD6	16	525
56	Q9KZD6	Q9KZD6	16	525
57	Q9KZD6	Q9KZD6	16	525
58	Q9KZD6	Q9KZD6	16	525
59	Q9KZD6	Q9KZD6	16	525
60	Q9KZD6	Q9KZD6	16	525
61	Q9KZD6	Q9KZD6	16	525
62	Q9KZD6	Q9KZD6	16	525
63	Q9KZD6	Q9KZD6	16	525
64	Q9KZD6	Q9KZD6	16	525
65	Q9KZD6	Q9KZD6	16	525
66	Q9KZD6	Q9KZD6	16	525
67	Q9KZD6	Q9KZD6	16	525
68	Q9KZD6	Q9KZD6	16	525
69	Q9KZD6	Q9KZD6	16	525
70	Q9KZD6	Q9KZD6	16	525
71	Q9KZD6	Q9KZD6	16	525
72	Q9KZD6	Q9KZD6	16	525
73	Q9KZD			

Order	Species	Count	Percentage
1	Q9KZD6	16	36.3
2	Q59995	16	36.3
3	Q949Y3	16	36.3
4	Q949Y3	16	36.3
5	Q949Y3	16	36.3
6	Q949Y3	16	36.3
7	Q949Y3	16	36.3
8	Q949Y3	16	36.3
9	Q949Y3	16	36.3
10	Q949Y3	16	36.3
11	Q949Y3	16	36.3
12	Q949Y3	16	36.3
13	Q949Y3	16	36.3
14	Q949Y3	16	36.3
15	Q949Y3	16	36.3
16	Q949Y3	16	36.3
17	Q949Y3	16	36.3
18	Q949Y3	16	36.3
19	Q949Y3	16	36.3
20	Q949Y3	16	36.3
21	Q949Y3	16	36.3
22	Q949Y3	16	36.3
23	Q949Y3	16	36.3
24	Q949Y3	16	36.3
25	Q949Y3	16	36.3
26	Q949Y3	16	36.3
27	Q949Y3	16	36.3
28	Q949Y3	16	36.3
29	Q949Y3	16	36.3
30	Q949Y3	16	36.3
31	Q949Y3	16	36.3
32	Q949Y3	16	36.3
33	Q949Y3	16	36.3
34	Q949Y3	16	36.3
35	Q949Y3	16	36.3
36	Q949Y3	16	36.3
37	Q949Y3	16	36.3
38	Q949Y3	16	36.3
39	Q949Y3	16	36.3
40	Q949Y3	16	36.3
41	Q949Y3	16	36.3
42	Q949Y3	16	36.3
43	Q949Y3	16	36.3
44	Q949Y3	16	36.3
45	Q949Y3	16	36.3
46	Q949Y3	16	36.3
47	Q949Y3	16	36.3
48	Q949Y3	16	36.3
49	Q949Y3	16	36.3
50	Q949Y3	16	36.3
51	Q949Y3	16	36.3
52	Q949Y3	16	36.3
53	Q949Y3	16	36.3
54	Q949Y3	16	36.3
55	Q949Y3	16	36.3
56	Q949Y3	16	36.3
57	Q949Y3	16	36.3
58	Q949Y3	16	36.3
59	Q949Y3	16	36.3
60	Q949Y3	16	36.3
61	Q949Y3	16	36.3
62	Q949Y3	16	36.3
63	Q949Y3	16	36.3
64	Q949Y3	16	36.3
65	Q949Y3	16	36.3
66	Q949Y3	16	36.3
67	Q949Y3	16	36.3
68	Q949Y3	16	36.3
69	Q949Y3	16	36.3
70	Q949Y3	16	36.3
71	Q949Y3	16	36.3
72	Q949Y3	16	36.3
73	Q949Y3	16	36.3
74	Q949Y3	16	36.3
75	Q949Y3	16	36.3
76	Q949Y3	16	36.3
77	Q949Y3	16	36.3
78	Q949Y3	16	36.3
79	Q949Y3	16	36.3
80	Q949Y3	16	36.3
81	Q949Y3	16	36.3
82	Q949Y3	16	36.3
83	Q949Y3	16	36.3
84	Q949Y3	16	36.3
85	Q949Y3	16	36.3
86	Q949Y3	16	36.3
87	Q949Y3	16	36.3
88	Q949Y3	16	36.3
89	Q949Y3	16	36.3
90	Q949Y3	16	36.3

Grain	Grain yield (t/ha)	Grain yield (t/ha)	Grain yield (t/ha)
4	38.5	Q59995	Q94973
43	57.5	Q59995	Q91d99
42	47.5	Q59995	Q91d99
5	47.5	Q59995	Q91d99

17: sp_alcney.
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed.
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	45	61.6	326	10	Q9ZT80	Q9ZT80 hordeum vul
2	58.9	283	5	Q9VZK0		Q9VZK0 dirosophila
3	58.9	323	5	Q9S5V2		Q9S5V2 dirosophila
4	43	381.9	525	16	Q9KZD6	Q9KZD6 streptomyces
5	43	381.5	525	16	Q9S995	Q9S995 synecocyst
6	42	57.5	424	16	Q949Y3	Q949Y3 aridiposist
7	42	57.5	475	10	Q949Y3	Q949Y3 gryza sativ
8	41	56.2	590	16	Q9LD99	Q9LD99 delinococcus
9	41	56.2	224	10	Q9RU88	Q9RU88 apis mellif
10	41	56.2	1998	5	Q95YM8	Q95YM8 dirosophila
11	40	54.8	302	5	Q95RE3	Q95RE3 dirosophila
12	40	54.8	452	5	Q9VIT1	Q9VIT1 dirosophila
13	40	54.8	766	16	Q9A0E7	Q9A0E7 staphaena s
14	40	54.8	1266	16	Q8YWG7	Q8YWG7 sus scrofa
15	39	53.4	135	6	Q9Z049	Q9Z049 dirosophila
16	39	53.4	179	5	Q9VPZ0	Q9VPZ0 dirosophila
17	39	53.4	222	4	Q9UNH34	Q9UNH34 homo sapien

17	39	53.4	251	13	08UP3	091008 streptococcus
18	39	53.4	308	18	091008	091008 xenopus
19	39	53.4	353	13	09W702	08xx9 dirosophila
20	39	53.4	401	5	08SX9	09xh7 streptomyc
21	39	53.4	506	16	09ZB7	09nq7 homo sapien
22	39	53.4	507	4	09N07	09b27 homo sapien
23	39	53.4	507	10	09PVT7	091y7 arabidopsis
24	39	53.4	649	5	09PVT7	02627 dictyostell
25	39	53.4	872	5	026257	043161 homo sapien
26	39	53.4	3926	10	043161	09up5 homo sapien
27	39	53.4	3851	5	09UP5	09f93 oryza sativ
28	38.5	52.7	123	10	09FW03	09m5 bradyrhizob
29	38.5	52.7	3310	2	09AM03	09n65 campylobact
30	38	52.1	75	16	09PNC6	09n66 campylobact
31	38	52.1	176	16	09BD04	028d4 rhizobium
32	38	52.1	271	5	022543	02543 cenorhabdi
33	38	52.1	287	17	08TTC7	08t4c methanosarc
34	38	52.1	293	16	08PFC30	08f30 brucella me
35	38	52.1	322	5	09S520	095520 dirosophila
36	38	52.1	628	10	08O977	08O977 arabidopsis
37	38	52.1	628	10	08O977	08O979 arabidopsis
38	38	52.1	630	10	P93027	P93027 arabidops
39	38	52.1	856	4	08IEE1	08tel1 homo sapien
40	38	52.1	3942	11	08B737	08B737 mus musculu
41	37.5	51.4	786	5	09R2L4	09r24 deinococcus
42	37	50.7	221	5	020281	020281 caenorhabdi
43	37	50.7	221	5	08SM14	08SM14 encephalito
44	37	50.7	222	3	09P8R1	09P8r1 sclerotinia
45	37	50.7	255	16	08XP77	08XP77 clostridium

ALIGNMENTS

RESULT 1			PRT;	326 AA.
09ZTWO	PRELIMINARY;			
ID 09ZTWO				
AC 01-MAY-1999 (TREMBlrel. 10, Created)				
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)				
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)				
DT ABA-responsive protein.				
DS Hordeum vulgare (Barley).				
OS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;				
OC Eukaryota; Viridiplantae; Liliopsida; Poales; Poaceae; Pooidae;				
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;				
OC Trilicaceae; Hordeum.				
OC NCBI_TaxID=4513;				
OX 11				
RN SEQUENCE FROM N. A.				
RP Liu J.-H., Luo W., Cheng K.-J., Mohapatra S.S., Hill R.D.;				
RT Identification and characterization of a novel barley gene that is				
RT ABA-inducible and expressed specifically in embryo and aleurone."				
RT J. Exp. Bot. 50:727-728(1999).				
DR EMBL: AF026538; AAD09343.1; "				
DR InterPro: IPR004182; GRAM_dom.				
DR Pfam: PF02893; GRAM_1				
DR SEQ 326 AA; 34301 MW; 5A6653196BD6AADE CRC64;				
Query Match	61.6%;	Score 45;	DB 10;	Length 326;
Best Local Similarity	70.0%;	Pred. NO. 5.9;		
Matches 7; Conservative		1; Indels	0;	Gaps
4 TEGRYTAPOP 13				
:11:11111				
DB 76 TEGRYTAPOP 85				
RESULT 2				
09VZRO	PRELIMINARY;		PRT;	283 AA.
ID 09VZRO				
AC 01-MAY-2000 (TREMBlrel. 13, Created)				
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)				

01-JUN-2000 (TREMBlrel. 14, Last annotation update)
CG12016.
Drosophila melanogaster (Fruit fly).
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
NCBI_Taxid=7227;
[1]
SEQUENCE FROM N.A.
STRAIN-BERKELEY;
MEDLINE=20196006; PubMed=10731132;
Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
Amaratunga P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
Wan K.H., Doyle C., Rogers Y.-H.C., Blazer V., Chai M., Peltier B.D.,
Abra J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Balaban D.,
Ballew R.M., Basu P., Baxendale J., Bayraktaroglu L., Beasley E.M.,
Beeson K.Y., Benos P.V., Bernier B.P., Bhandari D., Bolshakov S.,
Borova K.C., Botchan M.R., Bouck B., Butler H., Cantley L.C.,
Burris K.C., Busam D.A., Butler H., Cantley L.C., Center A., Chandra I.,
Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
de Paolo S., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
Dodson K.J., Evans J., Fink J., Garg N.S., Gelber R., Gelfand E.,
Giles C., Gabrielian A.E., Gong P., Gottlieb J.H., Guan P., Harris M.,
Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
Hoslin D., Houston K.A., Howland T.J., Mei M.-H., Ikegami C.,
Jalil M.B., Kalish J., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
Lasko P., Lei J., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
Mount S.M., Moshina N.V., Mobarry C., Morris J., Moshirei A.,
Neelson S.M., Nelson K.A., Nixon K., Nusskern D.R., Paeb J.M.,
Palazzo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
Shue B.C., Siden-Kimms I., Simpson M., Skupski M.P., Smith T.,
Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
Syllas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissenbach J.,
Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao O., Zheng L.,
Zheng X.H., Zhong P.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
"The genome sequence of Drosophila melanogaster."
Science 287:2185-2195(2000).
EMBL: AEO03477; AF47759.1;
DR FlyBase: FBgn0035436; CG12016.
SO SEQUENCE 283 AA; 33018 MW; FF397060D499BC90 CRC64;
Query Match
Best Local Similarity 58.9%; Score 43; DB 5; Length 283;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
QY 4 IEGRYTAPQPO 14
DB 93 IKGRHVAPEPO 103
RESULT 3
ID 095SV2 PRELIMINARY; PRT; 323 AA.
AC 095SV2;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DE SD05789P. (TREMBlrel. 20, Last annotation update)
GN CG12016.
OS Drosophila melanogaster (Fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophillidae; Drosophila.
NCBI_Taxid=7227;
[1]
SEQUENCE FROM N.A.
Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
Gonzalez M., Guerin H., Li P., Liao G., Miranda A., Mungall C.J.,
RA Munoz J., Pacio J., Paragas V., Park S., Phouanavong S., Wan K.,
RA Yu C., Lewis S.E., Rubin G.M., Celniker S.,
RL Submitted (09-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY060471; AL25510.1;
DR FlyBase: FBgn0035436; CG12016.
SO SEQUENCE 323 AA; 37869 MW; 780E1FC22C297DF2 CRC64;
Query Match
Best Local Similarity 58.9%; Score 43; DB 5; Length 323;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
QY 4 IEGRYTAPQPO 14
DB 93 IKGRHVAPEPO 103
RESULT 4
ID 09KZD6 PRELIMINARY; PRT; 525 AA.
AC 09KZD6;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DE Hypothetical protein SC06953.
GN SC06953 OR SC67.06C.
OS Streptomyces coelicolor.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycetaceae; Streptomyces.
NCBI_Taxid=1902;
[1]
SEQUENCE FROM N.A.
STRAIN-A3(2)/M45;
Bentley S.D., Chater K.F., Cerdano-Parraga A.-M., Challis G.L.,
RA Thompson N.R., James K.D., Harris D.E., Quail M.A., Kleser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth M.,
RA Huang C.H., Kleser T., Larke L., Murphy L., Oliver K., O'Neill S.,
RA Rabbinowitsch E., Rajandream M.A., Rutherford K., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Squires S., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J., Barrett B.G., Parhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
coelicolor A3(2)."
RL Nature 417:141-147(2002).
DR EMBL: AL353870; CAB85014.1;
KW Hypothetical protein.
SO SEQUENCE 525 AA; 57338 MW; 4A7BE97C8CAEB24 CRC64;
Query Match
Best Local Similarity 72.7%; Score 43; DB 16; Length 525;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 3 STEGRYAPQPO 13
DB 451 SIRGLFTAPQPO 461
RESULT 5
ID 059995 PRELIMINARY; PRT; 424 AA.
AC 059995;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DE 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
GN CG12016.
OS Drosophila melanogaster (Fruit fly).

Tang C., Toriumi M., Yamamura Y., Yu G., Yu S., Ishida L., Jones T.,
 Carninci P., Chen H., Cheuk R., Hayashizaki Y., Bowser L.,
 Kamita A., Karlin-Neumann G., Kawai J., Kim C., Koeseema E., Lam B.,
 Lin J., Meyers M.C., Miranda M., Narusaka M., Nguyen M., Palm C.J.,
 Sakurai T., Satou M., Seki M., Shinn P., Southwick A., Tracy S.E.,
 Shinozaki K., Davis R.W., Ecker J.R., Theologis A.,
 "Full length cDNA of gene TE15.10/At5g34850."
 Submitted (Aug-2001) to the EMBL/GenBank/DBJ databases.

[2]
 SEQUENCE FROM N.A.
 Yamada K., Banb J., Chan M.M., Chang C.H., Chang E., Dale J.M.,
 Dang J.M., Goldsmith A.D., Lee J.W., Onodera C.S., Quach H.L.,
 Tang C., Toriumi M., Wu H.C., Yamamura Y., Yu G., Bowser L.,
 Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,
 Kamita A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,
 Meyers M.C., Miranda M., Narusaka M., Nguyen M., Palm C.J.,
 Sakurai T., Satou M., Seki M., Shinn P., Southwick A., Shinozaki K.,
 Davis R.W., Ecker J.R., Theologis A.,
 "Arabidopsis Open Reading Frame (ORF) Clones."
 Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
 -1- CATALYTIC ACTIVITY: AN ORTHOPHOSPHORIC MONOESTER + H(2O) = AN
 ALCOHOL + PHOSPHATE.
 CC EMBL: AY050812; AAK92747.1; -;
 DR EMBL: AY091415; AAM14354.1; -;
 DR InterPro: IPR004961; FN.II.
 DR InterPro: IPR004843; M-peptidase.
 DR InterPro: IPR004844; S/T_phosphatase.
 DR Pfam: PF00149; Metallophos; 1.
 DR SMART: SM00060; FNC3; 1.
 KW Hydrolase; Iron; Zinc.
 DR SEQUENCE 475 AA; 55009 MW; 7C73A161EE16327E CRC64;

OY 1 QASTGRTAPAP 13
 1 : 11 : 111
 Db 391 QEGINGRTPEPP 403

RESULT 7
 ID Q9LID99 PRELIMINARY; PRT; 590 AA.
 AC Q9LID99;
 DT 01-OCT-2000 (TEMBLrel. 15, Created)
 DT 01-OCT-2000 (TEMBLrel. 15, Last sequence update)
 DT 01-JUN-2001 (TEMBLrel. 17, Last annotation update)
 DE Hypothetical protein.
 OS Oryza sativa (Rice).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Eriactoidae; Oryzaceae; Oryza.
 OX NCBI_TaxID=4530;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. NIPPONBARE;
 RA Sasaki T., Matsumoto T., Yamamoto K.;
 RT "Oryza sativa nippobare(GA3) genomic DNA, chromosome 1, PAC
 clone;P0469E05.";
 RT Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. NIPPONBARE;
 RA Sasaki T., Matsumoto T., Yamamoto K.;
 RT "Oryza sativa nippobare(GA3) genomic DNA, chromosome 1, PAC
 clone;P0511C01.";
 RT Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF002480; BAA95849.1; -;
 DR EMBL: AF002070; BAA95849.1; -;
 DR InterPro: IPR003662; sub_transporter.
 DR PROSITE: PS00216; SUGAR_TRANSPORT_1; UNKNOWN_1.
 DR SEQUENCE 590 AA; 63418 MW; D6DEB9FA0CA165E CRC64;

Query Match
 Best Local Similarity 57.5%; Score 42; DB 10; Length 590;
 Matches 8; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
 QY 1 QASIEGRYAPQP 13
 DB 535 QRETEGRYAPQP 547

RESULT 8
 ID Q9R088 PRELIMINARY; PRT; 224 AA.
 AC Q9R088;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DE 01-JUN-2001 (TREMBlrel. 13, Last annotation update)
 GN MADH dehydrogenase I, C subunit.
 OS Deinococcus radiodurans.
 OC Bacteria; Thermus/Deinococcus group; Deinococci; Deinococcales;
 OC Deinococcaceae; Deinococcus.
 RX NCBI_TaxID=1299;
 [1]
 SEQUENCE FROM N.A.
 RC STRAIN=RL;
 RX MEDLINE=20036896; PubMed=10567266;
 RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
 RA Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
 RA Morfitt K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
 RA Vamathevan J.J., Lam P., McDonald L., Utterback T., Zaleski C.,
 RA Kachurav K.A., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
 RA Fraser C.M., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
 RT "Genome sequence of the radioresistant bacterium Deinococcus
 RT radiodurans R1."
 RL Science 286:1571-1577(1999).
 DR EMBL: AEO01994; AAF11071.1;
 DR TRIP: DR1504;
 DR InterPro: IPR001268; Complex1_30K.
 DR Pfam: PF00329; Complex1_30Kd_1.
 DR ProDom: PD001581; Complex1_30K; 1.
 DR PROSITE: PS00542; COMPLEX1_30K; 1.
 KW Complete proteome.
 SO SEQUENCE 224 AA; 25032 MW; 05CC3BDF11281B35 CRC64;
 Query Match
 Best Local Similarity 56.2%; Score 41; DB 16; Length 224;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 7 RYTAPOPO 14
 DB 80 RYTAPOPO 87

RESULT 9
 ID Q95YM8 PRELIMINARY; PRT; 1598 AA.
 AC Q95YM8;
 DT 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DE Mblk-1 protein.
 GN MBLK-1.
 OS Apis mellifera (Honeybee).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Hymenoptera; Apocrita; Aculeata;
 OC Apoidea; Apidae; Apis.
 GN NCBI_TaxID=7460;
 [1]
 SEQUENCE FROM N.A.
 RA Takeuchi H., Kage E., Sawata M., Kamikouchi A., Ohashi K., Ohara M.,
 RA Fujiyuki T., Kunieda T., Sekimizu K., Natori S., Kubo T.;

RT "Identification of a novel gene, Mblk-1, that encodes a putative
 RT transcription factor expressed preferentially in the large-type Kenyon
 RT cells of the honey bee brain."
 RL Insect Mol. Biol. 10:487-494(2001).
 DR EMBL: AB047034; BAB64310.1;
 SO SEQUENCE 1598 AA; 174929 MW; E5475BDD3ACB1EEF CRC64;

Query Match
 Best Local Similarity 56.2%; Score 41; DB 5; Length 1598;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 6 GRYPAPQP 13
 DB 995 GRYPAPQP 1002

RESULT 10
 ID Q95RE3 PRELIMINARY; PRT; 302 AA.
 AC Q95RE3;
 DT 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DE 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 GN CG10263.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 GN NCBI_TaxID=7227;
 [1]
 SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY;
 RA Stapleton M., Brokstein P., Hong L., Agbayan A., Carlson J.,
 RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
 RA Gonzalez M., Guarini H., Li P., Liao G., Miranda A., Mungall C.J.,
 RA Nuncio J., Pacieb J., Paragas V., Park S., Phoumenavong S., Wan K.,
 RA Yu C., Lewis S.E., Rubin G.M., Celinker S.,
 RL Submitted (OCT-2001) to the EMBL/Genbank/DBJ databases.
 CC -1- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.
 DR EMBL: AY061441; AAL28989.1;
 DR FlyBase: FBgn0032812; CG10263.
 DR InterPro: IPR001841; ZNF_ring.
 DR Pfam: PF00097; ZF-CRHC4; 1.
 DR PROSITE: PS00518; ZF_RING_1; UNKNOWN_1.
 KW Zinc-finger.
 SO SEQUENCE 302 AA; 32559 MW; 7A54C93A1FF32BC CRC64;

Query Match
 Best Local Similarity 54.8%; Score 40; DB 5; Length 302;
 Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
 QY 1 QASIEGRYAPQP 13
 DB 248 QAHINHRVAPQP 260

RESULT 11
 ID Q9VIT1 PRELIMINARY; PRT; 452 AA.
 AC Q9VIT1;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DE CG10263 protein.
 GN CG10263.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 GN NCBI_TaxID=7227;
 [1]
 SEQUENCE FROM N.A.

RC STRAIN-BERKELEY;
 MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Ceiniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amandides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Mortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazey R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA April J.F., Apathy A., An H.-J., Andrews-Pfankoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhattacharya P., Bottier P.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
 RA Butts K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahle C., Davenport L.B., Davies P.,
 RA De Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Donnes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Fostler C., Gabriellian A.E., Garg N.S., Gelbalt W.M., Glasser K.,
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Idegami C.,
 RA Jajani M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshireit A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tecor C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yen R.-F., Zaveli J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
 RT "The genome sequence of *Drosophila melanogaster*."
 RL Science 287:2185-2195(2000).
 CC -1- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.
 DR EMBL: AF003663; AAF53834.1; -
 DR Flybase: FBgn0032812; CG10263.
 DR InterPro: IPR001841; Znf_ring.
 DR Pfam: PF00097; ZF-C3HC4; 1.
 DR SMART: SM00184; RING; 1.
 DR PROSITE: PS00518; ZF_RING_1; 1.
 KW zinc-finger.
 SQ SEQUENCE 452 AA; 47001 MW; 9763885BE66BFA69 CRC64;
 QY 1 QASIEGRYAPQP 13
 DB 248 QAHINRHVAPQP 260

Query Match 54.8%; Score 40; DB 5; Length 452;
 Best Local Similarity 61.5%; Pred. No. 70;
 Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

RESULT 12
 Q9A0E7 PRELIMINARY: PRT; 736 AA.
 AC Q9A0E7;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE Hypothetical protein SPY0807.
 OS Streptococcus pyogenes.
 OC Bacteria; Firmicutes; Clostridium group; Lactobacillales;
 OC Streptococcaceae; Streptococcus.
 NCBI_TaxID=1314;
 RN [1]

RP SEQUENCE FROM N.A.
 RC STRAIN-SF370 / ATCC 700294 / SEROTYPE M1;
 RA MEDLINE=21192684; PubMed=11296296;
 RA Ferretti J.J., McShan W.M., Ajdic D.J., Savic D.J., Savic G., Lyon K.,
 RA Primeaux C., Sezate S., Suvorov A.N., Kenton S., Lai H.S., Lin S.P.,
 RA Qian Y., Jia H.G., Najaf F.Z., Ren Q., Zhu H., Song L., White J.,
 RA Yuan X., Clifton S.W., Roe B.A., McLaughlin R.;
 RT "Complete genome sequence of an M1 strain of *Streptococcus pyogenes*."
 RL Proc. Natl. Acad. Sci. U.S.A. 98:4658-4663(2001).
 DR EMBL: AE006531; AK33744.1; -
 DR InterPro: IPR00917; Sulfatase.
 DR Pfam: PF00884; Sulfatase; 1.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 736 AA; 82914 MW; A3228642886858EA CRC64;
 QY 4 IEGRYAPQP 14
 DB 249 VKGVAAPQP 259

Query Match 54.8%; Score 40; DB 16; Length 736;
 Best Local Similarity 54.5%; Pred. No. 12e+02;
 Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

RESULT 13
 Q8YWG7 PRELIMINARY: PRT; 1263 AA.
 AC Q8YWG7;
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Heterocyst glycolipid synthase.
 GN HcLE OR A11646.
 OS Anabaena sp. (Strain PCC 7120).
 OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
 NCBI_TaxID=103690;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=21595285; PubMed=11759840;
 RA Kaneko T., Nakamura Y., Molk C.P., Kuritz T., Sasamoto S.,
 RA Katsube A., Iriguchi M., Ishikawa A., Kawashima K., Kimura T.,
 RA Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
 RA Nakazaki N., Shimpo S., Sugimoto M., Takazawa M., Yamada M.,
 RA Yasuda M., Tabata S.;
 RT "Complete genomic sequence of the filamentous nitrogen-fixing
 cyanobacterium *Anabaena* sp. strain PCC 7120."
 RL DNA Res. 8:205-213(2001).
 DR EMBL: AP003586; BAB78012.1; -
 DR InterPro: IPR001227; Ac_transferase.
 DR InterPro: IPR00794; ketoacyl-synt.
 DR Pfam: PF00698; ketoacyl-synt.
 DR Pfam: PF00109; ketoacyl-synt; 1.
 DR Pfam: PF02801; ketoacyl-synt; 1.
 DR Pfam: PF00550; pp-binding; 1.
 DR Pfam: PF00507; ACP_DOMAIN; 1.
 DR PROSITE: PS00606; B_KETOACYL-SYNTHASE; UNKNOWN_1.
 KW Complete proteome.
 SQ SEQUENCE 1263 AA; 137716 MW; FBA1BFA4A08B1430 CRC64;
 QY 4 IEGRYAPQP 14
 DB 1154 LEPRYSAPQP 1164

Query Match 54.8%; Score 40; DB 16; Length 1263;
 Best Local Similarity 63.6%; Pred. No. 2.3e+02;
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

RESULT 14
 Q29049 PRELIMINARY: PRT; 135 AA.
 AC Q29049;
 DT 029049;
 NCBI_TaxID=1314;
 RN [1]

```

DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Heterogeneous nuclear ribonucleoprotein (Fragment).
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-SKELETAL MUSCLE;
RX MEDLINE=99307156; PubMed=10375634;
RA Davoli R., Zambonielli P., Bigli D., Fontanesi L., Russo V.;
RA "Analysis of expressed sequence tags of porcine skeletal muscle.";
RL Gene 233:181-188(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE-SKELETAL MUSCLE;
RA Davoli R., Fontanesi L., Cepica S., Musilova P., Stratil A.,
RA Rubens J.;
RA "The porcine poly(rC)-binding protein 2 (PCBP2) gene maps to
RT chromosome 5.";
RL AnIm. Genet. 30:161-168(1999).
DR EMBL; X94253; CNA63936.1; -.
DR HSSP; Q07244; 1KHM.
DR InterPro; IPR004087; KH_dom.
DR InterPro; IPR004088; KH_type_1.
DR Pfam; PF00013; KH-domain; 1.
DR SMART; SM00322; KH; 1.
DR PROSITE; PS50084; KH-type_1; 1.
KM Nucleocapsid; Ribonucleoprotein.
FT NON_TER
SQ SEQUENCE 135 AA; 14144 MW; 8789F8CF3D3F2FC CRC64;

Query Match 53.4%; Score 39; DB 6; Length 135;
Best Local Similarity 54.5%; Pred. No. 26;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 3 SIEGRYTAPOP 13
Db 14 TIQGYAIRPOP 24

RESULT 15
Q9VPZ0
ID 09VPZ0 PRELIMINARY; PRT: 179 AA.
AC 09VPZ0;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DE 01-MAY-2000 (TREMBlrel. 13, Last annotation update)
DE CG5139 protein.
GN CG5139
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
RA Burks K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,

```

```

RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laoko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Mekulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
DR EMBL; AE003587; AAF51394.1; -.
DR FlyBase; FBgn0031323; CG5139.
SQ SEQUENCE 179 AA; 21086 MW; E842E509CC016E8 CRC64;

Query Match 53.4%; Score 39; DB 5; Length 179;
Best Local Similarity 53.8%; Pred. No. 36;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 1 QASIEGRYTAPOP 13
Db 51 EAEIQTATPNP 63

```

Search completed: January 13, 2003, 09:59:48
 Job time : 56.1282 secs

Copyright © 1993 - 2003
using sw mod/250 Million cell updates/second
in time 12.2051 seconds
each core - 2003
5.1.5

262574

Scoring	Match	08
Minimum	Match	100%
Maximum	Match	45 summaries

Listing 111
Search: AA:*
Parents-AA:*
/1/iaa/5A-COMB.pep:*
/1/iaa/5A-COMB.pep:*

```

Issued: Pa6/prodata/1/aa/5B.COMB.pep:*
1: /cgn2_6/prodata/1/aa/6A.COMB.pep:*
2: /cgn2_6/prodata/1/aa/6B.COMB.pep:*
3: /cgn2_6/prodata/1/aa/PCrUS.COMB.pep:*
4: /cgn2_6/prodata/1/aa/backfill1.pep:*
5: /cgn2_6/prodata/1/aa/backfill2.pep:*
6: /cgn2_6/prodata/1/aa/backfill3.pep:*

```

Pred. No. is the number of the total score greater than or equal to and is derived by analysis

SUMMARIES

Serial Number	Description	Amount	Balance
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100

Result No.	Score	Query Match	length	DB	ID
1	45	61	6	321	US-09-417-286-4
2	46	40	3	229	US-08-028-483-2
3	36	49	3	229	US-08-401-815-2
4	35	47	9	229	US-08-415-591-2
5	35	47	9	229	US-09-145-944-6
6	35	47	9	229	US-09-103-887-7
7	35	47	9	229	US-09-282-36
8	35	47	9	229	US-08-555-713-4
9	35	47	9	229	US-08-583-562B-8
10	35	47	9	229	US-08-779-113-8
11	35	47	9	229	US-08-591-989-5
12	35	47	9	229	US-08-591-989-5
13	35	47	9	229	US-08-690-473-2
14	35	47	9	229	US-08-959-821A-2
15	35	47	9	229	US-09-543-659-2
16	35	47	9	229	US-08-583-659-2
17	35	47	9	229	US-08-583-659-2
18	35	47	9	229	US-08-583-659-2
19	35	47	9	229	US-08-583-659-2
20	35	47	9	229	US-08-583-659-2
21	35	47	9	229	US-08-583-659-2
22	35	47	9	229	US-08-583-659-2
23	35	47	9	229	US-08-583-659-2
24	35	47	9	229	US-08-583-659-2
25	35	47	9	229	US-08-583-659-2
26	35	47	9	229	US-08-583-659-2
27	35	47	9	229	US-08-583-659-2

28	33	45	2	302	4	US-09-037	143-47
29	33	45	2	302	4	US-09-049	-691-47
30	33	45	2	302	4	US-08-269	-174-47
31	33	45	2	302	4	US-08-286	-128A-47
32	33	45	2	302	4	US-09-332	346-47
33	33	45	2	302	4	US-09-332	192-47
34	33	45	2	302	4	US-09-037	
35	33	45	2	388	4	5240849-3	-593A-8
36	33	45	2	424	4	US-09-332	-011A-4
37	33	45	2	536	4	US-07-826	-509A-12
38	33	45	2	536	4	US-08-426	-509A-13
39	33	45	2	536	4	US-08-426	-509A-14
40	33	45	2	536	4	US-08-426	-509A-15
41	33	45	2	536	4	PCT-US95-05008	-11
42	33	45	2	537	4	PCT-US95-05008	-11
43	33	45	2	543	4	PCT-US95-05008	-14
44	33	45	2	543	4	PCT-US95-05008	-14
45	33	45	2	543	4	PCT-US95-05008	-14

[illegible]

ALIGNMENT

RESULT 1
US-09/417-286-4 Application US/09417286
Sequence 4, 6016656
PATENT INFORMATION:
APPLICANT: Liu, Jin-Hao
APPLICANT: Cheng, Ku-Juan
APPLICANT: Chen, Tei-Plant
INVENTION: 08919-019001
TITLE OF INVENTION: 08919-019001
FILE REFERENCE NUMBER: 1999-1013
4 0

```

CURRENT FILING DATE:      8
CURRENT FILING ID NOS.:   0
NUMBER OF FASTSEQ FOR WINDOWS VERSION 4.0 SOFTWARE:    4
SOFTWARE:                  4
SEO ID NO:                 321
SEO LENGTH:               321

```

TYPE:	PRT	Hordeum vulgare	Score 45;	DB 4;	Leu...	0;	Caps
ORGANISM:			61.6%;				
IS-09-417-286-4			70.0%;				
Query Match	Similarity						
Best Local	7;	Conservative					
Matches							
OY	4	IEGRYTAPOP 13					
		: 11: 1111					
Db	75	IEGRYTAPOP 84					

[illegible]

Mon Jan 13 10:35:11 2003

us-09-554-941-11.ra1

Page 2

ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY disk
OPERATING SYSTEM: PC-DOS/MS-DOS
CURRENT APPLICATION: Patent Release #1.0, Version #1.25
APPLICATION NUMBER: 435
FILING DATE: 09-MAR-1993
PRIORITY DATE: 09-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: MISTROCK, S. LESLIE
REGISTRATION NUMBER: 18,872
TELEPHONE: 212 790-9090
TELEFAX: 212 790-9090
SEQUENCE CHAR SET ID NO: 1
LENGTH: 229 amino acids
TYPE: amino acid
STRANDEDNESS: single
MOLECULE TYPE: linear
ORIGINAL SOURCE: peptide
ORGANISM: Rhodococcus rhodochrous
STRAIN: J-1 (FERM BP-1478)
US-08-028-463-2

Query Match
Best Local Similarity 49.3%; Score 36; DB 1; Length 229;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 1 QASIEGRTYRPOP 13
DB 99 QETIEGRTYRPOP 111

US-09-145-391-2
Sequence 2; Application US/09145391
Patent No. 6194171
GENERAL INFORMATION:
APPLICANT: Puist, Stefan M.
TITLE OF INVENTION: Nucleic Acids Encoding Ataxin-2 Binding Proteins,
FILE REFERENCE: CE 3063
CURRENT APPLICATION NUMBER: US/09/145,391
NUMBER OF SEQ ID NOS: 2
SOFTWARE: Patent In Ver. 2.0
SEQ ID NO 2 Patent In Ver. 2.0
LENGTH: 330
TYPE: PPT
ORGANISM: Homo sapiens
US-09-145-391-2

APPLICATION NUMBER: US 08/028,463
FILING DATE: 09-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: MISTROCK, S. LESLIE
REGISTRATION NUMBER: 18,872
TELEPHONE: 212 790-9090
TELEFAX: 212 790-9090
SEQUENCE CHAR SET ID NO: 1
LENGTH: 229 amino acids
TYPE: amino acid
STRANDEDNESS: single
MOLECULE TYPE: linear
ORIGINAL SOURCE: peptide
ORGANISM: Rhodococcus rhodochrous
STRAIN: J-1 (FERM BP-1478)
US-08-028-463-2

Query Match
Best Local Similarity 49.3%; Score 36; DB 1; Length 229;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 1 QASIEGRTYRPOP 13
DB 99 QETIEGRTYRPOP 111

US-09-145-391-2
Sequence 2; Application US/09145391
Patent No. 6194171
GENERAL INFORMATION:
APPLICANT: Puist, Stefan M.
TITLE OF INVENTION: Nucleic Acids Encoding Ataxin-2 Binding Proteins,
FILE REFERENCE: CE 3063
CURRENT APPLICATION NUMBER: US/09/145,391
NUMBER OF SEQ ID NOS: 2
SOFTWARE: Patent In Ver. 2.0
SEQ ID NO 2 Patent In Ver. 2.0
LENGTH: 330
TYPE: PPT
ORGANISM: Homo sapiens
US-09-145-391-2

Query Match
Best Local Similarity 49.3%; Score 36; DB 1; Length 330;
Matches 7; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 1 QASIEGRTYRPOP 13
DB 34 QNGIDPATYRPOP 46

US-09-133-944-6
Sequence 6; Application US/09133944
Patent No. 6280937
GENERAL INFORMATION:
APPLICANT: Luo, Ying
TITLE OF INVENTION: SHUTTL VECTORS
FILE REFERENCE: A66522/DJB/DV
CURRENT APPLICATION NUMBER: US/09/133,944
FILING DATE: 1999-08-14

US-09-133-944-6
Sequence 6; Application US/09133944
Patent No. 6280937
GENERAL INFORMATION:
APPLICANT: Luo, Ying
TITLE OF INVENTION: SHUTTL VECTORS
FILE REFERENCE: A66522/DJB/DV
CURRENT APPLICATION NUMBER: US/09/133,944
FILING DATE: 1999-08-14

US-09-133-944-6
Sequence 6; Application US/09133944
Patent No. 6280937
GENERAL INFORMATION:
APPLICANT: Luo, Ying
TITLE OF INVENTION: SHUTTL VECTORS
FILE REFERENCE: A66522/DJB/DV
CURRENT APPLICATION NUMBER: US/09/133,944
FILING DATE: 1999-08-14

Mon Jan 13 10:35:11 2003

us-09-554-941-11.ra1

EARLIER APPLICATION NUMBER: 09/133,949
 EARLIER FILING DATE: 1998-08-14
 NUMBER OF SEQ ID NOS: 39
 SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO: 6
 LENGTH: 70
 TYPE: PRT
 ORGANISM: Unknown
 FEATURE:
 OTHER INFORMATION: Description of Unknown Organism: minbody
 OTHER INFORMATION: presentation structure
 US-09-133-944-6

Query Match 47.9%; Score 35; DB 4; Length 70;
 Best Local Similarity 75.0%; Pred. No. 28;
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 2 ASIEGRT 9
 11:1111
 Db 43 ASYKGRYT 50

RESULT 6
 US-09-208-827-7
 Sequence 7, Application US/09208827
 Patent No. 6391582
 GENERAL INFORMATION:
 APPLICANT: Luo, Ying
 APPLICANT: Yu, Pei Wen
 APPLICANT: Lotens, James
 TITLE OF INVENTION: SHUTTLE VECTORS
 FILE REFERENCE: A66252-1/DJB/DAV
 CURRENT APPLICATION NUMBER: US/09/208,827
 CURRENT FILING DATE: 1998-12-09
 EARLIER APPLICATION NUMBER: 09/133,949
 EARLIER FILING DATE: 1998-08-14
 NUMBER OF SEQ ID NOS: 43
 SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO: 7
 LENGTH: 70
 TYPE: PRT
 ORGANISM: Unknown
 FEATURE:
 OTHER INFORMATION: Description of Unknown Organism: minbody
 OTHER INFORMATION: presentation structure
 US-09-208-827-7

Query Match 47.9%; Score 35; DB 4; Length 70;
 Best Local Similarity 75.0%; Pred. No. 28;
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 2 ASIEGRT 9
 11:1111
 Db 43 ASYKGRYT 50

RESULT 7
 551582-36
 Patent No. 5514582
 APPLICANT: CAPON, DANIEL J.; LASKY, LAURENCE A.
 TITLE OF INVENTION: RECOMBINANT DNA ENCODING HYBRID
 IMMUNOGLOBULINS
 NUMBER OF SEQUENCES: 43
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/185,670
 FILING DATE: 21-JAN-1994
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 986,931
 FILING DATE: 08-DEC-1992
 APPLICATION NUMBER: 808,122
 FILING DATE: 16-DEC-1991
 APPLICATION NUMBER: 440,625
 FILING DATE: 22-NOV-1989

APPLICATION NUMBER: 315,015
 FILING DATE: 23-FEB-1989
 SEQ ID NO: 36
 LENGTH: 120
 5514582-36

Query Match 47.9%; Score 35; DB 6; Length 120;
 Best Local Similarity 50.0%; Pred. No. 51;
 Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

OY 1 QASIEGRTAPOPO 14
 1:11:111111
 Db 44 QCTSEGVWTPVPQ 57

RESULT 8
 US-08-955-713-4
 Sequence 4, Application US/08955713
 Patent No. 5955308
 GENERAL INFORMATION:
 APPLICANT: SATHE, GANESH
 APPLICANT: MOONEY, JEFFREY
 APPLICANT: BERGSMAN, DEBK
 APPLICANT: HALSLEY, WENDY
 TITLE OF INVENTION: CDNA CLONE HEOAD54 THAT ENCODES
 NUMBER OF SEQUENCES: 4
 CORRESPONDENCE ADDRESSES:
 ADDRESSEE: RATNER & PRESTIA
 STREET: P. O. BOX 980
 CITY: VALLEY FORGE
 STATE: PA
 COUNTRY: USA
 ZIP: 19482
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: FastSeq for Windows Version 2.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/955,713
 FILING DATE: 23-OCT-1997
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 60/050,124
 FILING DATE: 18-JUN-1997
 ATTORNEY/AGENT INFORMATION:
 NAME: PRESTIA, PAUL F
 REGISTRATION NUMBER: 23,031
 REFERENCE/DOCKET NUMBER: GH-70087
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 610-407-0700
 TELEFAX: 610-407-0701
 TELEX: 846169
 INFORMATION FOR SEQ ID NO: 4:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 476 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-955-713-4

Query Match 47.9%; Score 35; DB 2; Length 476;
 Best Local Similarity 55.6%; Pred. No. 2,3e+02;
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 5 EGRTAPOP 13
 1:1:1111
 Db 356 QGCHTVPPQ 364

RESULT 9
 US-08-583-562B-8

```

Sequence 8, Application US/08583562B
Patent No. 5922570
GENERAL INFORMATION:
APPLICANT: Staunton, Donald
TITLE OF INVENTION: Cytoplasmic Modulators of Integrin
NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESSES:
ADDRESS: Marshall, O'Toole, Gerstein, Murray & Borun
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/583,562B
FILING DATE:
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Williams Jr., Joseph A.
REGISTRATION NUMBER: 38,659
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-474-6300
TELEFAX: 312-474-0448
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 2647 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-583-562B-8

Query Match
Best Local Similarity 47.9%; Score 35; DB 2; Length 2647;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 YTAPOP 13
Db 1712 YTAPOP 1717

RESULT 10
US-08-779-113-8
Sequence 8, Application US/08779113
Patent No. 5948891
GENERAL INFORMATION:
APPLICANT: Staunton, Donald E.
TITLE OF INVENTION: Cytoplasmic Modulators of Integrin
NUMBER OF SEQUENCES: 65
CORRESPONDENCE ADDRESSES:
ADDRESS: Marshall, O'Toole, Gerstein, Murray & Borun
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/779,113
FILING DATE:

```

```

CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Greta E. No. 5948891and
REGISTRATION NUMBER: 35,302
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-474-6300
TELEFAX: 312-474-0448
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 2647 amino acids
TYPE: amino acid
TOPOLOGY: not relevant
US-08-779-113-8

Query Match
Best Local Similarity 47.9%; Score 35; DB 2; Length 2647;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 YTAPOP 13
Db 1712 YTAPOP 1717

RESULT 11
US-08-591-989-2
Sequence 2, Application US/08591989
Patent No. 5795721
GENERAL INFORMATION:
APPLICANT: Ross S. Rabin, Sumedha Jayasena
TITLE OF INVENTION: HIGH AFFINITY NUCLEIC
NUMBER OF SEQUENCES: 87
CORRESPONDENCE ADDRESSES:
ADDRESS: Swanson & Bratschun, L.L.C.
STREET: 8400 East Prentice Avenue, Suite #200
CITY: Englewood
STATE: Colorado
COUNTRY: USA
ZIP: 80111
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 1.40 MB
OPERATING SYSTEM: MS-DOS
SOFTWARE: WORD PERFECT 6.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/591,989
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Barry J. Swanson
REGISTRATION NUMBER: 33,215
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 793-3433
TELEFAX: (303) 793-3433
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 235
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-591-989-2

Query Match
Best Local Similarity 46.6%; Score 34; DB 1; Length 235;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 6 GRYTAPOP 14
Db 1712 YTAPOP 1717

```

Db 26 GRTAGOPR 34

RESULT 12

US-08-591-989-5

Sequence 5, Application US/08591989

Patent No. 5795721

GENERAL INFORMATION:

APPLICANT: Ross S. Rabin, Sumeetha Jayasena

APPLICANT: and Larry Gold

TITLE OF INVENTION: HIGH AFFINITY NUCLEIC

TITLE OF INVENTION: ACID LIGANDS OF ICP4

NUMBER OF SEQUENCES: 87

CORRESPONDENCE ADDRESS:

ADDRESSEE: Swanson & Bratschun, L.L.C.

STREET: 8400 East Prentice Avenue, Suite #200

CITY: Englewood

STATE: Colorado

COUNTRY: USA

ZIP: 80111

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.5 inch, 1.40 MB

MEDIUM TYPE: storage

COMPUTER: IBM COMPATIBLE

OPERATING SYSTEM: MS-DOS

SOFTWARE: WORD PERFECT 6.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/591,989

FILING DATE:

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Barry J. Swanson

REGISTRATION NUMBER: 33,215

REFERENCE/DOCKET NUMBER: NEX 49

TELECOMMUNICATION INFORMATION:

TELEPHONE: (303) 793-3333

TELEFAX: (303) 793-3433

INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:

LENGTH: 323

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

US-08-591-989-5

Query Match

Best Local Similarity 46.6%; Score 34; DB 1; Length 323;

Best Local Similarity 66.7%; Pred. No. 2.2e+02;

Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 6 GRTAPQPO 14

Db 114 GRTAGOPR 122

RESULT 13

US-08-690-473-2

Sequence 2, Application US/08690473

Patent No. 5876923

GENERAL INFORMATION:

APPLICANT: Leopardi, Rosario

APPLICANT: Roizman, Bernard

TITLE OF INVENTION: HERPES SIMPLEX VIRUS ICP4 AS AN

TITLE OF INVENTION: INHIBITOR OF APOPTOSIS

NUMBER OF SEQUENCES: 2

CORRESPONDENCE ADDRESS:

ADDRESSEE: Arnold, White & Durkee

STREET: P.O. Box 4433

CITY: Houston

STATE: Texas

COUNTRY: USA

ZIP: 77210

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/690,473

FILING DATE: 26-JUL-1996

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Highlander, Steven L.

REGISTRATION NUMBER: 37,642

REFERENCE/DOCKET NUMBER: ARCD:239

TELECOMMUNICATION INFORMATION:

TELEPHONE: 512/474-7577

TELEFAX: 512/474-7577

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 1298 amino acids

TYPE: amino acid

STRANDEDNESS:

TOPOLOGY: linear

US-08-690-473-2

Query Match

Best Local Similarity 46.6%; Score 34; DB 2; Length 1298;

Best Local Similarity 66.7%; Pred. No. 1e+03;

Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 6 GRTAPQPO 14

Db 281 GRTAGOPR 289

RESULT 14

US-09-259-821A-2

Sequence 2, Application US/09259821A

Patent No. 6210926

GENERAL INFORMATION:

APPLICANT: LEOPARDI, ROSARIO

APPLICANT: ROIZMAN, BERNARD

TITLE OF INVENTION: HERPES SIMPLEX VIRUS ICP4 IS AN INHIBITOR OF APOPTOSIS

FILE REFERENCE: ARCD:317

CURRENT APPLICATION NUMBER: US/09/259,821A

CURRENT FILING DATE: 1999-03-01

PRIOR APPLICATION NUMBER: 08/690,473

PRIOR FILING DATE: 1996-07-26

NUMBER OF SEQ ID NOS: 2

SOFTWARE: Patent Ver. 2.1

SEQ ID NO 2

LENGTH: 1298

TYPE: PRT

ORGANISM: HERPES VIRUS, TYPE 1

US-09-259-821A-2

Query Match

Best Local Similarity 46.6%; Score 34; DB 4; Length 1298;

Best Local Similarity 66.7%; Pred. No. 1e+03;

Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 6 GRTAPQPO 14

Db 281 GRTAGOPR 289

RESULT 15

US-08-843-659-2

Sequence 2, Application US/08843659

Patent No. 6218103

GENERAL INFORMATION:

APPLICANT: Leopardi, Rosario

APPLICANT: Roizman, Bernard

TITLE OF INVENTION: HERPES SIMPLEX VIRUS US3 AND ICP4 AS

TITLE OF INVENTION: INHIBITORS OF APOPTOSIS

NUMBER OF SEQUENCES: 6

CORRESPONDENCE ADDRESS:

ADDRESSEE: Arnold, White & Durkee

STREET: P.O. Box 4433
CITY: Houston
STATE: Texas
COUNTRY: United States
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/843,659
FILING DATE: Concurrently Herewith
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Highlander, Steven L.
REGISTRATION NUMBER: 37,642
REFERENCE/DOCKET NUMBER: ARSB:519
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 418-3000
TELEFAX: (512) 474-7577
INFORMATION FOR SEQ. ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1298 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
US-08-843-659-2

Query Match 46.6%; Score 34; DB 4; Length 1298;
Best Local Similarity 66.7%; Pred. No. 1e+03;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 6 GRYTAPQPG 14
Db 281 GRFTAGQPR 289

Search completed: January 13, 2003, 09:56:26
Job time: 14.2051 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 13, 2003, 09:53:15 ; Search time 14 Seconds
(without alignments)
41.476 Million cell updates/sec

Title: US-09-554-941-11
Perfect score: 73
Sequence: 1 QASISGRTAPQPO 14

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	42	57.5	1	092H58 streptomyces
2	39	53.4	1	061990 mus musculus
3	39	53.4	1	015366 homo sapien
4	39	53.4	1	099217 homo sapien
5	39	53.4	1	091144 mus musculus
6	39	53.4	1	P15278 drosophila
7	38	52.1	1	P07636 bacterioph
8	37	50.7	1	P21819 manduca sex
9	37	50.7	1	001102 mus musculus
10	37	50.7	1	P98106 rattus norv
11	37	50.7	1	LEM3_RAT
12	37	50.7	1	LEM3_SHEEP
13	37	50.7	1	HMPB_DROME
14	37	50.7	1	YPT3_CABEL
15	36	49.3	1	PURQ_MYCTU
16	36	49.3	1	NHBL_RHORI
17	36	49.3	1	VRAP_ECOLI
18	36	49.3	1	A2BP_MOUSE
19	36	49.3	1	A2BP_HUMAN
20	36	49.3	1	YOB1_CABEL
21	36	49.3	1	ORAI1_MOUSE
22	36	49.3	1	SRG1_XENLA
23	36	49.3	1	SRG2_XENLA
24	36	49.3	1	FNX_XIPHE
25	36	49.3	1	PRKA_STRCO
26	36	49.3	1	GRPE_BACST
27	35.5	48.6	1	PMBA_ECOLI
28	35.5	48.6	1	GAOA_DACDE
29	35.5	48.6	1	KFLA_MOUSE
30	35	47.9	1	PCXB_ACICA
31	35	47.9	1	YLAB_ECOLI
32	35	47.9	1	KIME_AERPE
33	35	47.9	1	DIV_ECOLI

Result	Score	Query Match Length	ID	Description
34	35	47.9	1	CAP1_DICDI
35	35	47.9	1	L756_CABEL
36	35	47.9	1	RHSE_ECOLI
37	35	47.9	1	XYND_RUMFL
38	35	47.9	1	Y104_SYNY3
39	35	47.9	1	GPC2_HUMAN
40	35	47.9	1	CR2_HUMAN
41	35	47.9	1	FNPT5_HUMAN
42	35	47.9	1	PLNA_HUMAN
43	34.5	47.3	1	ABG5_HUMAN
44	34.5	47.3	1	KFLA_HUMAN
45	34	46.6	1	CYCP_CHRVI

ALIGNMENTS

Result	Score	Query Match Length	ID	Description
1	42	57.5	1	092H58 streptomyces
2	39	53.4	1	061990 mus musculus
3	39	53.4	1	015366 homo sapien
4	39	53.4	1	099217 homo sapien
5	39	53.4	1	091144 mus musculus
6	39	53.4	1	P15278 drosophila
7	38	52.1	1	P07636 bacterioph
8	37	50.7	1	P21819 manduca sex
9	37	50.7	1	001102 mus musculus
10	37	50.7	1	P98106 rattus norv
11	37	50.7	1	LEM3_RAT
12	37	50.7	1	LEM3_SHEEP
13	37	50.7	1	HMPB_DROME
14	37	50.7	1	YPT3_CABEL
15	36	49.3	1	PURQ_MYCTU
16	36	49.3	1	NHBL_RHORI
17	36	49.3	1	VRAP_ECOLI
18	36	49.3	1	A2BP_MOUSE
19	36	49.3	1	A2BP_HUMAN
20	36	49.3	1	YOB1_CABEL
21	36	49.3	1	ORAI1_MOUSE
22	36	49.3	1	SRG1_XENLA
23	36	49.3	1	SRG2_XENLA
24	36	49.3	1	FNX_XIPHE
25	36	49.3	1	PRKA_STRCO
26	36	49.3	1	GRPE_BACST
27	35.5	48.6	1	PMBA_ECOLI
28	35.5	48.6	1	GAOA_DACDE
29	35.5	48.6	1	KFLA_MOUSE
30	35	47.9	1	PCXB_ACICA
31	35	47.9	1	YLAB_ECOLI
32	35	47.9	1	KIME_AERPE
33	35	47.9	1	DIV_ECOLI

CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: AF071885; AAC70948.1; -
 DR EMBL: AF442143; CAC09994.1; -
 DR HSSP: P19245; 1TYF.
 DR MEROPS: S14.UPW; -
 DR InterPro: IPR001907; CLP_protease.
 DR Pfam: PF00574; CLP_protease.1.
 DR PRINTS: PR00127; CLP_protease.1.
 DR TIGRFAMs: TIGR00493; clp.1.
 DR PROSITE: PS00382; CLP_PROTEASE_HIS.1.
 DR PROSITE: PS00381; CLP_PROTEASE_SER.1.
 DR Hydrolase; serine protease; Multigene family; Complete proteome.
 FT ACT_SITE 114 114
 FT ACT_SITE 139 139 BY SIMILARITY.
 SQ SEQUENCE 218 AA; 24180 MW; 67082B89EAC49F2E CRC64;

Query Match 57.5%; Score 42; DB 1; Length 218;
 Best Local Similarity 50.0%; Pred. No. 2.1;
 Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;
 QY 1 QASIEGRYAPQ 14
 Db 2 RAASQGRYTPQAE 15

RESULT 2

PCB2_MOUSE STANDARD; PRT; 362 AA.
 AC 061950; 061883; 062042;
 DT 01-NOV-1997 (Rel. 35; Created)
 DT 01-NOV-1997 (Rel. 35; Last sequence update)
 DT 15-JUN-2002 (Rel. 41; Last annotation update)
 DE Poly(rlc)-binding protein 2 (Alpha-CP2) (Putative heterogeneous nuclear
 DE ribonucleoprotein X) (hnRNP X) (CTBP) (CBP).
 GN PCB2 OR HNRNPX OR HNRPX OR CBP.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
 RA MEDLINE=93376518; PubMed=8367306;
 RA Hahn K.B., Kim G., Turch C., Smale S.T.;
 RT "Isolation of a murine gene encoding a nucleic acid-binding protein
 RT with homology to hnRNP K.";
 RL Nucleic Acids Res. 21:3894-3894 (1993).
 RN [2]
 RP SEQUENCE FROM N.A. AND PARTIAL SROUNCE (ISOFORM 2).
 RA STRAIN=C57BL/6 X 129/Ola; TISSUE=Liver;
 RA MEDLINE=94268912; PubMed=8208614;
 RA Goller M., Funke B., Gehe-Becker C., Kroege B., Lottspeich F.,
 RA Horak I.;
 RT "Murine protein which binds preferentially to oligo-C-rich single-
 RT stranded nucleic acids.";
 RL Nucleic Acids Res. 22:1885-1889 (1994).
 RN [3]
 RP SEQUENCE FROM N.A. (ISOFORM 3).
 RA STRAIN=C57BL/6 X CBA; TISSUE=Liver;
 RA Horak I.;
 RT Submitted (May-1996) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RA Makeyev A.V., Liehaber S.A.;
 RT "Identification of two novel mammalian genes establishes a subfamily
 RT of KH-domain RNA-binding proteins.";
 RL Submitted (Feb-2000) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: SINGLE-STRANDED NUCLEIC ACID BINDING PROTEIN THAT
 CC BINDS PREFERENTIALLY TO OLIGO DC.
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- ALTERNATIVE PRODUCTS: 3 ISOFORMS; 1 (SHOWN HERE), 2 AND 3; ARE

CC PRODUCED BY ALTERNATIVE SPLICING.
 CC -1- SIMILARITY: CONTAINS 3 KH DOMAINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: L19661; AA03705.1; -
 DR EMBL: X75947; CA53346.1; -
 DR EMBL: X79982; CA66619.1; -
 DR EMBL: AF236845; AAK14059.1; -
 DR EMBL: AF236842; AAK14059.1; JOINED.
 DR EMBL: AF236843; AAK14059.1; JOINED.
 DR EMBL: AF236844; AAK14059.1; JOINED.
 DR HSSP: 007244; 1KHM.
 DR MGD: MGI:108202; Pcbp2.
 DR InterPro: IPR004087; KH_dom.
 DR InterPro: IPR004088; KH_type_1.
 DR Pfam: PF00013; KH_domain.3.
 DR SMART: SM00322; KH.3.
 DR PROSITE: PS50084; KH_type_1.3.
 DR Nucleic protein; RNA-binding; Ribonucleoprotein; DNA-binding;
 DR Repeat; Alternative splicing.
 FT DOMAIN 13 75 KH 1.
 FT DOMAIN 97 162 KH 2.
 FT DOMAIN 194 224 KH 3.
 FT VASPLIC 194 224 MISSING (IN ISOFORM 2).
 FT VASPLIC 263 275 MISSING (IN ISOFORM 3).
 SQ SEQUENCE 362 AA; 38221 MW; 70C8A710E3BF3C0 CRC64;

Query Match 53.4%; Score 39; DB 1; Length 362;
 Best Local Similarity 54.5%; Pred. No. 13;
 Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
 QY 3 SIGGRYAPQ 13
 Db 227 TIQGYAIPQ 237

RESULT 3

PCB2_HUMAN STANDARD; PRT; 365 AA.
 AC 015366;
 DT 30-MAY-2000 (Rel. 39; Created)
 DT 30-MAY-2000 (Rel. 39; Last sequence update)
 DT 16-OCT-2001 (Rel. 40; Last annotation update)
 DE Poly(rlc)-binding protein 2 (Alpha-CP2) (hnRNP-E2).
 GN PCB2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=95331278; PubMed=7607214;
 RA Leffers H., Deigaard K., Celis J.E.;
 RT "Characterisation of two major cellular poly(rlc)-binding human
 RT proteins, each containing three K-homologous (KH) domains.";
 RL Eur. J. Biochem. 230:447-453 (1995).
 CC -1- FUNCTION: MAJOR CELLULAR POLY(RC)-BINDING PROTEIN. BINDS ALSO
 CC POLY(RU).
 CC -1- SUBCELLULAR LOCATION: LOOSELY BOUND IN THE NUCLEUS. MAY SHUTTLE
 CC BETWEEN THE NUCLEUS AND THE CYTOSOL.
 CC -1- TISSUE SPECIFICITY: DETECTED IN ALL TISSUES EXAMINED.
 CC -1- PTM: PHOSPHORYLATED. THE NON-PHOSPHORYLATED FORM(S) EXHIBITED THE
 CC STRONGEST POLY(RC)-BINDING ACTIVITY.
 CC -1- SIMILARITY: CONTAINS 3 KH DOMAINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL; X78136; CAA55015.1; -
DR HSSP; Q07244; 1KHM; PCBP2.
DR Genew; HGNC:8648; PCBP2.
DR MIM; 601210; -
DR InterPro: IPR004087; KH_dom.
DR InterPro: IPR004088; KH_type-1.
DR Pfam; PF00013; KH-domain; 3.
DR SMART; SM00322; KH_3.
DR PROSITE; PS50084; KH_type_1; 3.
DR Nuclear protein; RNA-binding; Ribonucleoprotein; DNA-binding;
KW Phosphorylation; Repeat.
FT DOMAIN 13 75 KH 1.
FT DOMAIN 97 162 KH 2.
FT DOMAIN 287 351 KH 3.
FT SEQUENCE 365 AA; 38580 MW; 43F03D76FDC2C63 CRC64;
SQ

Query Match 53.4%; Score 39; DB 1; Length 365;
Best Local Similarity 54.5%; Pred. No. 13; Indels 0; Gaps 0;
Matches 6; Conservative 3; Mismatches 0;

QY 3 SIEGRTAPOP 13
Db 231 TIOGVAIPQP 241

RESULT 4
WDRC_HUMAN STANDARD; PRT; 423 AA.
ID WDRC_HUMAN Q96ZL7; Q9N718; Q96H00; Q9N780;
AC Q96ZL7; Q9N718; Q96H00; Q9N780;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE WD-repeat protein 12 (YTM1 homolog).
GN WDR12.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Matsumoto S.;
RT "Human homolog of Saccharomyces cerevisiae YTM1."
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
[2]
RP SEQUENCE FROM N.A.
RA Iseki T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,
RA Magatsuna M., Hosoi T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
RA Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,
RA Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,
RA Yamamoto J., Wakamatsu A., Nakamura Y., Nagahari K., Masuno Y.,
RA Niimiya K., Iwayanagi T.;
RT "NEDD human cDNA sequencing project."
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
[3]
RP SEQUENCE FROM N.A.
RA TISSUE=Skin;
RC Strausberg R.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
-i- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAIN).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL; AF242546; AAF60355.1; -
DR EMBL; AF242546; AAF60355.1; -
DR EMBL; AK001743; BAA91875.1; -
DR EMBL; AK022781; BAB14242.1; -
DR EMBL; AK022782; BAB14243.1; -
DR EMBL; BC008082; AAB08082.1; -
DR Genew; HGNC:14098; WDR12.
DR InterPro: IPR001680; WD40.
DR Pfam; PF00400; WD40; 7.
DR PRINTS; PR00320; GPROTEINRPT.
DR PRODOM; PD000018; WD40; 7.
DR SMART; SM00678; WD_REPEATS_1; 2.
DR PROSITE; PS50082; WD_REPEATS_2; 5.
DR PROSITE; PS50294; WD_REPEATS_REGION; 1.
KW Repeat; WD repeat; Polymorphism.
FT REPEAT 99 137 WD 1.
FT REPEAT 138 180 WD 2.
FT REPEAT 187 226 WD 3.
FT REPEAT 225 293 WD 4.
FT REPEAT 295 334 WD 5.
FT REPEAT 340 360 WD 6.
FT REPEAT 384 422 WD 7.
FT VARIANT 72 72 M -> V.
FT VARIANT 89 89 /FTID=VAR_012863.
FT VARIANT 89 89 E -> G.
FT VARIANT 286 286 /FTID=VAR_012864.
FT VARIANT 286 286 E -> G.
FT CONFLICT 75 75 I -> V (IN REF. 1).
FT CONFLICT 333 333 L -> M (IN REF. 1).
FT SEQUENCE 423 AA; 47708 MW; 6D088C640AC981D8 CRC64;
SQ

Query Match 53.4%; Score 39; DB 1; Length 423;
Best Local Similarity 75.0%; Pred. No. 15; Indels 0; Gaps 0;
Matches 6; Conservative 2; Mismatches 0;

QY 7 RYTAPOPO 14
Db 88 KYTAPOPE 95

RESULT 5
WDRC_MOUSE STANDARD; PRT; 423 AA.
ID WDRC_MOUSE Q9JUA4; Q9JUF5; Q9CSP3;
AC Q9JUA4; Q9JUF5; Q9CSP3;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE WD-repeat protein 12 (YTM1 homolog).
GN WDR12.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA TISSUE=Thymus;
RC MEDLINE=2168157; PubMed=11827460;
RA Nal B., Mohr E., Da Silva M.I., Tagett R., Navarro C., Carroll P.,
RA Depetris D., Vertuhy C., Jordan B.R., Ferrer P.;
RT "Wdr12, a mouse gene encoding a novel WD-repeat protein with a
RT notchless-like amino-terminal domain."
RL Genomics 79:77-86 (2002).
[2]
RP SEQUENCE FROM N.A.
RA STRAIN=129/SV X C57BL/6J;
RC Matsumoto S.;
RL "Mouse homologue of Saccharomyces cerevisiae YTM1."
Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
[3]

Query	7	RYTAPQPO 14	53.4%	Score 39	DB 1	Length 423
Db	88	KYTAPOPE 95	75.0% <td>Pred. No. 15<td></td><td></td></td>	Pred. No. 15 <td></td> <td></td>		
		Matches 6	Conservative 2	Mismatches 9	Indels 0	Gaps 0

ID	FASTA DROME	STANDARD	PRT	508 AA
AC	P15278.03VJ89;			
DT	01-APR-1990 (Rel. 14, Created)			
DT	01-APR-1990 (Rel. 14, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Fasciclin III precursor (FAS III).			
GN	FAS3 OR CG5803			
OS	Drosophila melanogaster (Fruit fly).			
OC	Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;			
OC	Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;			
OC	Muscomorpha; Ephydroidea; Drosophilidae; Drosophillidae; Drosophila.			
OX	NCBI_taxid=7221;			
RP	SEQUENCE FROM N.A.			
RP	[1]			
RX	MEDLINE=90030406; PubMed=2509076;			
RA	Snow P.M., Bieber A.J., Goodman C.S.;			
RA	"Fasciclin III: a novel homophilic adhesion molecule in Drosophila.";			
RT	Cell 59:313-323(1989).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RP	STRAIN=Berkley;			
RC	MEDLINE=20196006; PubMed=10731132;			
RX	Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,			
RA	Amaratides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,			
RA	George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,			
RA	Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,			
RA	Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,			
RA	Abbil J.F., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,			
RA	Ballev R.M., Baau A.A., Baxendale J., Bayraktaroglu L., Beasley E.M.,			
RA	Beecon K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,			
RA	Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,			
RA	Burtis K.C., Busam D.A., Butler J., Cadieu E., Center A., Chandra I.,			
RA	Cherry J.M., Cvely S., Dahlke C., Davenport R., Davies P.,			
RA	de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,			
RA	Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,			
RA	Durbin K.J., Evangelista C.C., Ferraz C., Ferlita S., Fleischmann W.,			
RA	Foster C., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,			
RA	Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,			
RA	Hoslin D., Houston K.A., Howard T.J., Wei M.-H., Ibegwam C.,			
RA	Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,			
RA	Laslo P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,			
RA	Liu X., Matel B., McIntosh T.C., McLeod M.P., McPherson D.,			
RA	Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,			
RA	Mout S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,			
RA	Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,			
RA	Palazzo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,			
RA	Rehert K., Remington K., Saunders R.D.C., Scheefer F., Shen H.,			
RA	Shue B.C., Sidenklamos I., Simpson M., Skupski M.P., Smith T.,			
RA	Spieler E., Spradling A.C., Stapleton M., Strong R., Sun E.,			
RA	Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,			
RA	Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,			
RA	Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao O.A.,			
RA	Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,			
RA	Gibbs R.A., Zhong E.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,			
RT	"The genome sequence of Drosophila melanogaster.";			
RL	Science 287:2185-2195(2000).			

PLAYS A ROLE IN AXON OUTGROWTH, GUIDANCE AND FASCICULATION OF THE DEVELOPING NERVOUS SYSTEM.
 CC DEVELOPING NERVOUS SYSTEM.
 CC SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- TISSUE SPECIFICITY: EXPRESSED ON DIFFERENT SUBSETS OF AXON BUNDLES (FASCICLES) IN INSECT EMBRYOS.
 CC -1- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY.
 CC -1- SIMILARITY: CONTAINS 2 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
 CC -1- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE V-TYPE DOMAIN.
 CC
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).
 CC
 CC EMBL: M27813; AAA28532.1; -
 CC EMBL: AE003658; AAF53655.1; -
 CC PIR: A33378; A33378
 CC FlyBase: FBgn0000636; Fasn3.
 CC Cell adhesion; Glycoprotein; Repeat; Immunoglobulin domain;
 CC Transmembrane; Signal; Neurogenesis; Phosphorylation.
 CC
 CC SIGNAL 1 20
 CC CHAIN 21 508 FASCICLIN III.
 CC DOMAIN 21 346 EXTRACELLULAR (POTENTIAL).
 CC TRANSMEM 347 370 POTENTIAL.
 CC DOMAIN 371 508 CYTOPLASMIC (POTENTIAL).
 CC DOMAIN 44 106 IG-LIKE V-TYPE DOMAIN.
 CC DOMAIN 143 218 IG-LIKE C2-TYPE DOMAIN 1.
 CC DOMAIN 256 310 IG-LIKE C2-TYPE DOMAIN 2.
 CC MOD_RES 21 21 PYRROLIDONE CARBOXYLIC ACID (PROBABLE).
 CC DISULFID 150 211 POTENTIAL.
 CC CARBOHYD 62 62 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC CARBOHYD 160 160 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC CARBOHYD 257 257 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC CARBOHYD 300 300 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC CARBOHYD 382 382 PHOSPHORYLATION (POTENTIAL).
 CC MOD_RES 459 459 PHOSPHORYLATION (POTENTIAL).
 CC CONFLICT 1 15 MSRIYFICLAAILFD -> MEKQAEISGYAMH (IN REF.
 CC
 CC SEQUENCE 508 AA; 55883 MW; 6E39EA0580697DAF C6C64;
 CC
 CC Query Match 53.4%; Score 39; DB 1; Length 508;
 CC Best Local Similarity 50.0%; Pred. No. 18;
 CC Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
 CC
 CC 1 QASIEGRYAPQ 14
 CC 1 :||:|:|:|
 CC 279 QGRDGRYSAYEPQ 292
 CC
 CC RESULT 7
 CC ID TRA_BPMU STANDARD: PRT: 663 AA.
 CC AC P07636; P06021; (Created)
 CC DT 01-APR-1988 (Rel. 07, Last sequence update)
 CC DT 01-FEB-1996 (Rel. 33, Last annotation update)
 CC DT 16-OCT-2001 (Rel. 40, Last annotation update)
 CC DE Transposase.
 CC GN A OR 3.
 CC OS Bacteriophage Mu.
 CC OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Myoviridae;
 CC OC Mu-like viruses.
 CC OC NCBI_TaxID=10677;
 CC [1]
 CC RP SEQUENCE FROM N.A.
 CC MEDLINE=86067968; PubMed=2999776;
 CC Hershby R.M., Getzoff E.D., Baldwin D.L., Miller J.L., Chaconas G.;
 CC "Primary structure of phage mu transposase: homology to mu
 CC repressor";
 CC RT Proc. Natl. Acad. Sci. U.S.A. 82:7676-7680(1985).
 CC RN [2]

SEQUENCE FROM N.A.
 RA Press H., Brauer B., Schmidt C., Kamp D.;
 RT "Sequence of the left end of Mu."; (eds.);
 RL (in) Symonds N., Toussaint A., van de putte P., Howe M.M. (eds.);
 RL Phage Mu, pp.277-296, Cold Spring Harbor Laboratory Press,
 RL New York (1987).
 CC [3]
 CC RP SEQUENCE FROM N.A.
 CC RA Morgan G., Hatfull G., Hendrix R.;
 CC RT "Genome of bacteriophage Mu and comparison with the haemophilus
 CC influenzae Mu-like prophage Flumu.";
 CC Submitted (Aug-1998) to the EMBL/GenBank/DBJ databases.
 CC [4]
 CC RP SEQUENCE OF 1-88 FROM N.A.
 CC MEDLINE=83012203; PubMed=6214696;
 CC Press H., Kamp D., Kaumann R., Brauer B., Dellus H.;
 CC "Nucleotide sequence of the immunity region of bacteriophage Mu.";
 CC Mol. Gen. Genet. 186:315-321(1982).
 CC [5]
 CC RP SEQUENCE OF 1-84 FROM N.A.
 CC MEDLINE=83218562; PubMed=6222246;
 CC Toussaint A., Faellen M., Desmet L., Allet B.;
 CC "The products of gene A of the related phages Mu and D108 differ in
 CC their specificities."; Mol. Gen. Genet. 190:70-79(1983).
 CC [6]
 CC RP STRUCTURE BY NMR OF 1-76.
 CC MEDLINE=95187707; PubMed=7881904;
 CC Clubb R.T., Michalski J.G., Savilahti H., Mizuuchi K.,
 CC Gronenborn A.M., Clore G.M.;
 CC "A novel class of winged helix-turn-helix protein: the DNA-binding
 CC domain of Mu transposase."; Structure 2:1041-1048(1994).
 CC [7]
 CC RP STRUCTURE BY NMR OF 76-174.
 CC MEDLINE=98070329; PubMed=9405381;
 CC Schumacher S., Clubb R.T., Cai M., Mizuuchi K., Clore G.M.,
 CC Gronenborn A.M.;
 CC "Solution structure of the Mu end DNA-binding Ibeta subdomain of
 CC phage Mu transposase: modular DNA recognition by two tethered
 CC domains."; EMBO J. 16:7532-7541(1997).
 CC [8]
 CC RP STRUCTURE BY NMR OF 173-247.
 CC MEDLINE=98035037; PubMed=9367742;
 CC Clubb R.T., Schumacher S., Mizuuchi K., Gronenborn A.M., Clore G.M.;
 CC "Solution structure of the I gamma subdomain of the Mu end
 CC DNA-binding domain of phage Mu transposase."; J. Mol. Biol. 273:19-25(1997).
 CC [9]
 CC RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS) OF 248-574.
 CC MEDLINE=95354202; PubMed=7628012;
 CC Rice P., Mizuuchi K.;
 CC "Structure of the bacteriophage Mu transposase core: a common
 CC structural motif for DNA transposition and retroviral integration."; Cell 82:209-220(1995).
 CC -1- FUNCTION: THIS TRANSPOSASE IS ESSENTIAL FOR INTEGRATION,
 CC REPLICATION-TRANSPOSITION, AND EXCISION OF MU DNA.
 CC -1- MISCELLANEOUS: MU CAN TRANSPOSE ITS DNA INTO MULTIPLE SITES IN
 CC MANY BACTERIAL GENOMES AND MEDIANE A VARIETY OF DNA
 CC REARRANGEMENTS. TRANSPOSITION ENHANCER (ENCODED BY GENE B).
 CC BY GENE A) AND TRANSPOSITION ENHANCER (ENCODED BY GENE B).
 CC -1- MISCELLANEOUS: UNLIKE OTHER TRANSPOSONS MU HAS DISTINCTLY BINDS
 CC SEQUENCES AT ITS LEFT AND RIGHT ENDS. TRANSPOSASE APPARENTLY BINDS
 CC 3 SPECIFIC BLOCKS OF SEQUENCES AT EACH END OF MU DNA.
 CC -1- MISCELLANEOUS: THE A GENE IS REGULATED BY THE REPRESSOR C, WHICH
 CC BINDS TO AN OPERATOR SEQUENCE & TURNS OFF TRANSCRIPTION. REPRESSOR
 CC C CAN, AT HIGH CONCENTRATIONS, OCCUPY ALMOST THE EXACT SAME SITES
 CC ON MU ENDS AS THE TRANSPOSASE, AND TRANSPOSASE CAN BIND TO
 CC FRAGMENTS CONTAINING THE MU OPERATOR SEQUENCE.
 CC -1- SIMILARITY: STRONG. TO H. INFLUENZAE H11478.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

FT	CARBOHYD	398	398	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	603	603	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	654	654	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	661	661	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	679	679	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	LIPID	745	745	PALMITATE (BY SIMILARITY).	
FT	SITE	756	759	ENDOCYTOSIS SIGNAL (PROBABLE).	
FT	SITE	724	724	A -> E (IN REF. 2).	
SEQ	CONFLICT	768 AA;	83098 MW;	E5173074D2F66E68 CRC64;	
SEQ	SEQUENCE	768 AA;	83098 MW;	E5173074D2F66E68 CRC64;	
Query Match		50.7%;	Score 37;	DB 1;	Length 768;
Best Local Similarity		38.5%;	Prod. No. 65;		
Matches	5; Conservative	5;	Mismatches	3;	Indels 0; Gaps 0;
OY	1 OASIGRTYAPQF 13	:	:	:	:
Dd	553 ECTVSGRMSADPP 565	:	:	:	:
RESULT 10					
ID	LEM3_RAT	STANDARD;	PRT;	768 AA.	
AC	p98106;				
DT	01-FEB-1996 (Rel. 33, Created)				
DT	01-FEB-1996 (Rel. 33, Last sequence update)				
DT	01-JUN-2002 (Rel. 41, last annotation update)				
DE	P-selectin precursor (granule membrane protein 140) (GMP-140) (PADGEM)				
DE	(CD62P) (Leukocyte-endothelial cell adhesion molecule 3) (LECAM3).				
GN	Strp.				
OC	Rattus norvegicus (Rat).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.				
OX	NCBI_Taxid=10116;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=Lung;				
RX	MEDLINE=94333817; PubMed=7520013;				
RA	Anchamach J.A., Oliver M.G., Anderson D.C., Manning A.M.;				
RT	"Cloning, sequence comparison and in vivo expression of the gene				
RT	encoding rat P-selectin.";				
RL	Gene 145:251-255(1994).				
CC	- FUNCTION: CA(2+)-DEPENDENT RECEPTOR FOR MYELOID CELLS THAT BINDS				
CC	TO CARBOHYDRATES ON NEUTROPHILS AND MONOCYTES. MEDIATES THE				
CC	INTERACTION OF ACTIVATED ENDOTHELIAL CELLS OR PLATELETS WITH				
CC	LEUKOCYTES. THE LIGAND RECOGNIZED IS SIALYL-LEWIS X.				
CC	- SUBCELLULAR LOCATION: Type I membrane protein.				
CC	- TISSUE SPECIFICITY: EXPRESSED IN ALL TISSUES EXAMINED: SPLEEN,				
CC	LUNG, BRAIN, LIVER, HEART, KIDNEY, THYMUS, SMALL INTESTINE.				
CC	- INDUCTION: BY ACUTE INFLAMMATION (PROBABLE).				
CC	- SIMILARITY: BELONGS TO THE SELECTIN/LECTIN FAMILY.				
CC	- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.				
CC	- SIMILARITY: CONTRAINS 8 SUSHI (SCR) DOMAINS; RAT P-LECTIN LACKS THE				
CC	HUMAN SUSH1-2 EQUIVALENT.				
CC	-----				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/).				
CC	-- or send an email to license@isb-sib.ch . -----				
CC	EMBL: L23088; AAA60325.1; -				
DR	HSSP: L23088; IPFB.				
DR	InterPro: IPR000561; EGF-like.				
DR	InterPro: IPR001304; Lectin_C.				
DR	InterPro: IPR002396; Selectin.				
DR	InterPro: IPR000436; Sushi_SCR_CCP.				
DR	pfam: PF00059; Lectin_c_1.				
DR	pfam: PF00084; sushi_8.				
DR	PRINTS: PR00343; SELECTIN.				
DR	SMART: SM00032; CCP; 8.				

DE	P-selectin precursor (Granule membrane protein 140) (GMP-140) (PADGEM)
DE	(CD62p) (Leukocyte-endothelial cell adhesion molecule 3) (LECAM3).
GN	SELP.
OS	Ovis aries (Sheep).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC	Bovidae; Caprinae; Ovis.
OX	NCB1_TaxID=9940;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	TISSUE=Heart;
RA	Burns S.A., Neufeld E.J., Donady J.J.;
RL	Submitted (JUL-1994) to the EMBL/Genbank/DBJ databases.
CC	-1- FUNCTION: CA(2+)-DEPENDENT RECEPTOR FOR MYELOID CELLS THAT BINDS
CC	TO CARBOHYDRATES ON NEUTROPHILS AND MONOCYTES. MEDIATES THE
CC	INTERACTION OF ACTIVATED ENDOTHELIAL CELLS OR PLATELETS WITH
CC	LEUCOCYTES. THE LIGAND RECOGNIZED IS STALYI-LEMIS X.
CC	-1- SUBCELLULAR LOCATION: Type I membrane protein.
CC	-1- SIMILARITY: BELONGS TO THE SELECTIN/LECAM FAMILY.
CC	-1- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
CC	-1- SIMILARITY: CONTAINS 8 SUSHI (SCR) DOMAINS.
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC	the European Bioinformatics Institute. There are no restrictions on its
CC	use by non-profit institutions as long as its content is in no way
CC	modified and this statement is not removed. Usage by and for commercial
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC	or send an email to license@isb-sib.ch).
CC	-----
DR	EMBL: J34270; AA859261.1; -
DR	HSSP: P6109; IESB.
DR	InterPro: IPR000561; EGF-like.
DR	InterPro: IPR001304; Lectin.C.
DR	InterPro: IPR002396; Selectin.
DR	InterPro: IPR000436; Sush1_SCR_CCP.
DR	Pfam: PF00008; EGF_1.
DR	Pfam: PF00059; lectin_c; 1.
DR	Pfam: PF00084; sush1; 8.
DR	PRINTS: SM00343; SELECTIN.
DR	SMART: SM00332; CCP; 8.
DR	SMART: SM00334; CLECT; 1.
DR	SMART: SM00181; EGF; 1.
DR	PROSITE: PS00022; EGF_1; 1.
DR	PROSITE: PS01186; EGF_2; 1.
DR	PROSITE: PS00615; C-TYPE_LECTIN_1; 1.
DR	PROSITE: PS50041; C-TYPE_LECTIN_2; 1.
KW	Cell adhesion; Transmembrane; Glycoprotein; EGF-like domain; Lectin;
KW	Selectin; signal; Sush1; Repeat.
FT	STGNL 1 32
FT	CHAIN 1 32
FT	POTENTIAL.
FT	P-SELECTIN.
FT	DOMAIN 33 769
FT	TRANSMEM 33 717
FT	DOMAIN 718 734
FT	DOMAIN 735 769
FT	DOMAIN 769 799
FT	DOMAIN 799 829
FT	DOMAIN 829 859
FT	DOMAIN 859 889
FT	DOMAIN 889 919
FT	DOMAIN 919 949
FT	DOMAIN 949 979
FT	DOMAIN 979 1009
FT	DOMAIN 1009 1039
FT	DOMAIN 1039 1069
FT	DOMAIN 1069 1099
FT	DOMAIN 1099 1129
FT	DOMAIN 1129 1159
FT	DOMAIN 1159 1189
FT	DOMAIN 1189 1219
FT	DOMAIN 1219 1249
FT	DOMAIN 1249 1279
FT	DOMAIN 1279 1309
FT	DOMAIN 1309 1339
FT	DOMAIN 1339 1369
FT	DOMAIN 1369 1399
FT	DOMAIN 1399 1429
FT	DOMAIN 1429 1459
FT	DOMAIN 1459 1489
FT	DOMAIN 1489 1519
FT	DOMAIN 1519 1549
FT	DOMAIN 1549 1579
FT	DOMAIN 1579 1609
FT	DOMAIN 1609 1639
FT	DOMAIN 1639 1669
FT	DOMAIN 1669 1699
FT	DOMAIN 1699 1729
FT	DOMAIN 1729 1759
FT	DOMAIN 1759 1789
FT	DOMAIN 1789 1819
FT	DOMAIN 1819 1849
FT	DOMAIN 1849 1879
FT	DOMAIN 1879 1909
FT	DOMAIN 1909 1939
FT	DOMAIN 1939 1969
FT	DOMAIN 1969 1999
FT	DOMAIN 1999 2029
FT	DOMAIN 2029 2059
FT	DOMAIN 2059 2089
FT	DOMAIN 2089 2119
FT	DOMAIN 2119 2149
FT	DOMAIN 2149 2179
FT	DOMAIN 2179 2209
FT	DOMAIN 2209 2239
FT	DOMAIN 2239 2269
FT	DOMAIN 2269 2299
FT	DOMAIN 2299 2329
FT	DOMAIN 2329 2359
FT	DOMAIN 2359 2389
FT	DOMAIN 2389 2419
FT	DOMAIN 2419 2449
FT	DOMAIN 2449 2479
FT	DOMAIN 2479 2509
FT	DOMAIN 2509 2539
FT	DOMAIN 2539 2569
FT	DOMAIN 2569 2599
FT	DOMAIN 2599 2629
FT	DOMAIN 2629 2659
FT	DOMAIN 2659 2689
FT	DOMAIN 2689 2719
FT	DOMAIN 2719 2749
FT	DOMAIN 2749 2779
FT	DOMAIN 2779 2809
FT	DOMAIN 2809 2839
FT	DOMAIN 2839 2869
FT	DOMAIN 2869 2899
FT	DOMAIN 2899 2929
FT	DOMAIN 2929 2959
FT	DOMAIN 2959 2989
FT	DOMAIN 2989 3019
FT	DOMAIN 3019 3049
FT	DOMAIN 3049 3079
FT	DOMAIN 3079 3109
FT	DOMAIN 3109 3139
FT	DOMAIN 3139 3169
FT	DOMAIN 3169 3199
FT	DOMAIN 3199 3229
FT	DOMAIN 3229 3259

Query Match 50.7%; Score 37; DB 1; Length 769;
 Best Local Similarity 46.2%; Pred. No. 65;
 Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

FT DISULFID 292 319 BY SIMILARITY.
 FT DISULFID 324 368 BY SIMILARITY.
 FT DISULFID 354 381 BY SIMILARITY.
 FT DISULFID 386 430 BY SIMILARITY.
 FT DISULFID 416 443 BY SIMILARITY.
 FT DISULFID 448 492 BY SIMILARITY.
 FT DISULFID 478 505 BY SIMILARITY.
 FT DISULFID 510 554 BY SIMILARITY.
 FT DISULFID 540 567 BY SIMILARITY.
 FT DISULFID 581 625 BY SIMILARITY.
 FT DISULFID 611 638 BY SIMILARITY.
 FT DISULFID 643 687 BY SIMILARITY.
 FT DISULFID 673 700 BY SIMILARITY.
 FT CARBOHYD 54 54 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 80 80 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 180 180 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 212 212 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 219 219 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 347 347 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 398 398 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 604 604 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 655 655 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 662 662 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 680 680 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT SITE 757 760 ENDOCYTOSIS SIGNAL (PROBABLE).
 FT VARIANT 566 566 L -> T.
 FT VARIANT 579 579 S -> V.
 SQ SEQUENCE 769 AA; 84317 MW; 23E42575D60FAB15 CRC64;

RESULT 12
 HMPB_DROME STANDARD; PRT; 798 AA.
 ID HMPB_DROME
 AC P31264;
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Homeotic proboscipedia protein.
 GN PB.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
 OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
 OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
 OC NCBI_Taxid=7227;
 RN [1]
 RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
 RP STRAIN-Canton-S;
 RX MEDLINE=92224882; PubMed=1348688;
 RA Cribbs D.L., Pultz M.A., Johnson D., Mazulla M., Kaufman T.C.;
 RT "Structural complexity and evolutionary conservation of the
 RT Drosophila homeotic gene proboscipedia.";
 RL EMBL J. 11.1437-1449(1992).
 CC -1- FUNCTION: Sequence-specific transcription factor which is part of
 CC a developmental regulatory system that provides cells with
 CC specific positional identities on the anterior-posterior axis.
 CC Controls development of mouthparts, and labial and maxillary
 CC palps.
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- ALTERNATIVE PRODUCTS: At least 4 isoforms; 1 (shown here), 2, 3
 CC and 4; are produced by alternative splicing.
 CC -1- SIMILARITY: BELONGS TO THE ANTP HOMEOBOX FAMILY.
 CC PROBOSCIPEDIA SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

 DR EMBL: X63729; CAA45272.1;
 DR EMBL: X63728; CAA45271.1;
 DR PIR: S19852; S19852.
 DR PIR: S20881; S20881.
 DR HSSP: P02833; 9ANT.
 DR TRANSFAC: T02084;
 DR FlyBase: FBgn003040; pb.
 DR InterPro: IPR001827; Antennapedia.
 DR InterPro: IPR001356; Homeobox.
 DR Pfam: PF00046; homeobox.1.
 DR PRINTS: PR00025; ANTENNAPEDIA.
 DR PRINTS: PR00024; HOMEOBOX.
 DR ProDom: PD000010; Homeobox.1.
 DR SMART: SM00389; HOX.1.
 DR PROSITE: PS00027; HOMEOBOX_1; 1.
 DR PROSITE: PS00032; ANTENNAPEDIA; 1.
 DR PROSITE: PS00071; HOMEOBOX_2; 1.
 DR Homeobox; DNA-binding: Developmental protein; Nuclear protein;
 KW Alternative splicing.
 KW Homeobox; 169
 FT DOMAIN 164 169 ANTP-TYPE HEXAPEPTIDE.
 FT DNA_BIND 198 257 HOMEOBOX.
 FT DOMAIN 472 494 GLN-RICH.
 FT DOMAIN 535 580 HIS/GLN-RICH.
 FT VARSPIC 184 189 GNSIT -> A (IN ISOFORM 2).
 FT VARSPIC 189 194 TREYVE -> K (IN ISOFORM 3).
 FT VARSPIC 184 193 MISSING (IN ISOFORM 4).
 SQ SEQUENCE 798 AA; 85559 MW; 886289FA6C1758 CRC64;

Query Match 50.7%; Score 37; DB 1; Length 798;
 Best Local Similarity 38.5%; Pred. No. 68;
 Matches 5; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 1 QASIEGRYTAPOP 13
 Db 516 QQALDGEVLSPKP 528

RESULT 13
 YP73_CAEEL STANDARD; PRT; 4385 AA.
 ID YP73_CAEEL
 AC 009222;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 01-NOV-1995 (Rel. 32, Last annotation update)
 DE Hypothetical 486.0 kDa protein B0228.3 in chromosome II.
 GN B0228.3.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 OC Rhabditidae; Peloderae; Caenorhabditis.
 OC NCBI_Taxid=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RP STRAIN-Bristol N2;
 RA Lembuch D., Waterston R.;
 RL Submitted (Mar-1995) to the EMBL/GenBank/DBJ databases.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: U23168; AAC38807.1;
 CC WormRep: B0228.3; CE01744.
 CC Hypothetical protein.

SEQUENCE 4385 AA: 486003 MW: 50B871B6C45FA23 CRC64;

Query Match

Best Local Similarity 50.7%; Score 37; DB 1; Length 4385;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 3 SIEGRYAPQ 14

DB 1318 SIIGKINAPPEQ 1329

RESULT 14

PURQ.MYCTU

ID PURQ.MYCTU STANDARD: PRT: 224 AA.

AC P71841;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DE Phosphoribosylformylglycinamide synthase I (EC 6.3.5.3) (PGAM

GN synthase I)

OS PURQ OR RV0788 OR MT0813 OR MTCY369.32.

OC Mycobacterium tuberculosis.

OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;

OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.

NCBI_TaxID=1773;

[1]

SEQUENCE FROM N.A.

RA STRAIN-H37Rv;

RA MEDLINE=9825987; PubMed=9634230;

RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,

RA Gordon S.V., Eigmeier K., Gas S., Barry C.E. III, Tekaia F.,

RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,

RA Davies R., Devlin K., Feldwell T., Gentles S., Hamlin N., Holroyd S.,

RA Hornsby T., Jagsen K., Kroch A., McLean J., Moule S., Murphy L.,

RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,

RA Rutter S., Seeger K., Skellon S., Squares S., Squares R.,

RA Salston J.E., Taylor K., Whitehead S., Barrell B.G.;

RA "Deciphering the biology of Mycobacterium tuberculosis from the

RT complete genome sequence."

RL Nature 393:537-544(1998).

[2]

SEQUENCE FROM N.A.

RA STRAIN-CDC 1551 / Oshkosh;

RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,

RA Peterson J., Deboy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,

RA Kolonay J.F., Nelson W.C., Umayan L.A., Ermolaeva M.D., Salzberg S.L.,

RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,

RA Bishai W.;

RA "Whole genome comparison of Mycobacterium tuberculosis clinical and

RT laboratory strains."

RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.

CC -1- CATALYTIC ACTIVITY: ATP + N(2)-formyl-N(1)-(5-phospho-D-

CC ribosyl)glycinamide + L-glutamine + H(2)O = ADP + phosphate + 2-

CC (formamido)-N(1)-(5-phospho-D-riboseyl)acetamide + L-glutamate.

CC -1- PATHWAY: De novo purine biosynthesis, fourth step.

CC -1- SUBUNIT: HETERODIMER OF TWO SUBUNITS.

CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).

CC -1- SIMILARITY: CONTAINS 1 TYPE-1 GLUTAMINE AMIDOTRANSFERASE DOMAIN.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@sib-sib.ch).

DR EMBL: 280236; CA02378.1;

DR EMBL: AE006971; AKA5055.1;

DR TIGR: MT0813;

DR Tubercolin; RV0788;

DR InterPro: IPR000991; GATase_1.

DR Pfam: PF01965; ThiJ: 1

DR PROSITE: PS00442; GATASE_TYPE_1: 1.

KW Purine biosynthesis; Ligase; Glutamine amidotransferase;

FT ACYL_SITE 87

SO SEQUENCE 224 AA: 23633 MW: 46B37215D117DAE CRC64;

Query Match

Best Local Similarity 49.3%; Score 36; DB 1; Length 224;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 5 EGRYAPQ 12

DB 145 EGRYAPQ 152

RESULT 15

NHBL_RHQRH

ID NHBL_RHQRH STANDARD: PRT: 229 AA.

AC P21220;

DT 01-MAY-1991 (Rel. 18, Created)

DT 01-DEC-1992 (Rel. 24, Last sequence update)

DE 15-JUN-2002 (Rel. 41, Last annotation update)

DE High-molecular weight cobalt-containing nitrile hydratase subunit beta

GN (EC 4.2.1.84) (H-nitrilase) (H-NHase).

OS Rhodococcus rhodochrous.

OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;

OC Actinomycetales; Corynebacterineae; Nocardiaceae; Rhodococcus.

NCBI_TaxID=1829;

[1]

SEQUENCE FROM N.A. AND SEQUENCE OF 1-28 AND 69-78.

RA STRAIN-J1;

RA MEDLINE=92096459; PubMed=1840499;

RA Kobayashi M., Nishiyama M., Nagasawa T., Horinouchi S., Beppu T.,

RA Yamada H.;

RA "Cloning, nucleotide sequence and expression in Escherichia coli of

RT two cobalt-containing nitrile hydratase genes from Rhodococcus

RL rhodochrous J1."

RL Biochim. Biophys. Acta 1129:23-33(1991).

[2]

SEQUENCE FROM N.A.

RA STRAIN-J1;

RA MEDLINE=96210630; PubMed=8633053;

RA Komeda H., Kobayashi M., Shimizu S.;

RA "Characterization of the gene cluster of high-molecular-mass nitrile

RT hydratase (H-NHase) induced by its reaction product in Rhodococcus

RL rhodochrous J1."

RL Proc. Natl. Acad. Sci. U.S.A. 93:4267-4272(1996).

[3]

SEQUENCE OF 1-28.

RA STRAIN-J1;

RA MEDLINE=91192028; PubMed=2013281;

RA Nagasawa T., Takeuchi K., Yamada H.;

RA "Characterization of a new cobalt-containing nitrile hydratase

RT purified from urea-induced cells of Rhodococcus rhodochrous J1."

RL Eur. J. Biochem. 196:581-589(1991)

CC -1- FUNCTION: NHASE CATALYZES THE HYDRATION OF VARIOUS NITRILE

CC COMPOUNDS TO THE CORRESPONDING AMIDES. INDUSTRIAL PRODUCTION OF

CC -1- CATALYTIC ACTIVITY: An aliphatic amide + a nitrile + H(2)O.

CC -1- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA CHAIN.

CC -1- INDUCTION: BY COBALT AND UREA OR CYCLOHEXANECARBOXAMIDE.

CC -1- SIMILARITY: BELONGS TO THE NITRILE HYDRATASE SUBUNIT BETA FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@sib-sib.ch).

DR EMBL: Y64359; CAA45709.1; -
 DR EMBL: D67027; BAA1043.1; -
 DR PIR: S14772; S14772.
 DR PIR: S19713; S19713.
 DR HSSP: P13449; 2AH7.
 DR InterPro: IPR003168; NHase_beta.
 DR Pfam: PF02211; NHase_beta; 1.
 KW Lyase.
 FT CONFLICT 7 7 T->I (IN REF. 3).
 FT CONFLICT 26 26 MISSING (IN REF. 3).
 SQ SEQUENCE 229 AA; 26321 MW; 73270C13358D4671 CRC64;

Query Match 49.3%; Score 36; DB 1; Length 229;
 Best Local Similarity 53.8%; Pred. No. 27;
 Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 QASIEGRYTAPOP 13
 | :||||| :|
 Db 99 QETLEGRYTDKRP 111

Search completed: January 13, 2003, 09:54:09
 Job time : 16.3333 secs

Mon Jan 13 10:35:13 2003

US-09-554-941-2.rai

Page 2

```

RESULT 2
US-09-322-478-8
; Sequence 8, Application US/09322478
; Patent No. 6331662
; GENERAL INFORMATION:
; APPLICANT: Wright, David A.
; TITLE OF INVENTION: Plant Retroelements and Methods Related Thereto
; CURRENT APPLICATION NUMBER: US/09/322,478
; EARLIER FILING DATE: 1999-05-28
; EARLIER APPLICATION NUMBER: 60/087125
; SOFTWARE OF SEQ ID NOS: 41
; SEQ ID NO 8
; LENGTH: 532
; TYPE: PRT
; ORGANISM: Artificial Sequence
; OTHER INFORMATION: Description of Artificial Sequence: plant
US-09-322-478-8
Query Match
Best Local Similarity 63.2%; Score 36; DB 4; Length 532;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
OY 1 QASSTAPOP 10
DB 36 EGSSYSTSP 45

RESULT 3
US-09-322-478-18
; Sequence 18, Application US/09322478
; Patent No. 6331662
; GENERAL INFORMATION:
; APPLICANT: Wright, David A.
; TITLE OF INVENTION: Plant Retroelements and Methods Related Thereto
; CURRENT APPLICATION NUMBER: US/09/322,478
; EARLIER FILING DATE: 1999-05-28
; EARLIER APPLICATION NUMBER: 60/087125
; SOFTWARE OF SEQ ID NOS: 41
; SEQ ID NO 18
; LENGTH: 1802
; TYPE: PRT
; ORGANISM: Artificial Sequence
; OTHER INFORMATION: Description of Artificial Sequence: plant
US-09-322-478-18
Query Match
Best Local Similarity 60.0%; Score 36; DB 4; Length 1802;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
OY 1 QASSTAPOP 10
DB 36 EGSSYSTSP 45

RESULT 4
US-08-539-205A-6
; Sequence 6, Application US/08539205A
; Patent No. 6001619
; GENERAL INFORMATION:
; APPLICANT: Beach, David H.
; TITLE OF INVENTION: Plant Retroelements and Methods Related Thereto
; CURRENT APPLICATION NUMBER: US/08/539,205A
; EARLIER FILING DATE: 1999-05-28
; EARLIER APPLICATION NUMBER: 60/087125
; SOFTWARE OF SEQ ID NOS: 41
; SEQ ID NO 8
; LENGTH: 532
; TYPE: PRT
; ORGANISM: Artificial Sequence
; OTHER INFORMATION: Description of Artificial Sequence: plant
US-08-539-205A-6
Query Match
Best Local Similarity 61.4%; Score 35; DB 3; Length 834;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
OY 1 QASSTAPOP 11
DB 340 QPSYNSKPO 350
```

```

APPLICANT: Neisky, Bradley
TITLE OF INVENTION: Ubiquitin Ligases, and Uses Related Thereto
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
STREET: FOLEY, HOAG & ELIOT LLP
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109-2170
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: PC-DOS/MS-DOS
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/583,562B
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Williams Jr., Joseph A.
REGISTRATION NUMBER: 38,659
TELECOMMUNICATION INFORMATION: 27866/33033

US-08-583-562B-8
; Sequence 8, Application US/08583562B
; Patent No. 5922570
; GENERAL INFORMATION:
; APPLICANT: Staunton, Donald
; TITLE OF INVENTION: Cytoplasmic Modulators of Integrin
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; STREET: 233 South Wacker Drive, 6500 Sears Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: PC-DOS/MS-DOS
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/583,562B
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Williams Jr., Joseph A.
REGISTRATION NUMBER: 38,659
TELECOMMUNICATION INFORMATION: 27866/33033
```

TELEPHONE: 312-474-6300
TELEFAX: 312-474-0448
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 2647 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-583-562B-8

Query Match 61.4%; Score 35; DB 2; Length 2647;
Best Local Similarity 100.0%; Pred. NO. 1e+03; Indels 0; Gaps 0;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 YTAPOP 10
1712 YTAPOP 1717

RESULT 6
US-08-779-113-8
Sequence 8, Application US/08779113
Patent No. 5948891

GENERAL INFORMATION:
APPLICANT: Staunton, Donald E.
APPLICANT: Harris, Edith S.
TITLE OF INVENTION: Cytoplasmic Modulators of Integrin
TITLE OF INVENTION: Binding
NUMBER OF SEQUENCES: 65
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 233 South Wacker Drive, 6300 Sears Tower
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/779,113
FILING DATE:
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Greta E. No. 5948891and
REGISTRATION NUMBER: 35,302
REFERENCE/DOCKET NUMBER: 27866/33773
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-474-6300
TELEFAX: 312-474-0448
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 2647 amino acids
TYPE: amino acid
TOPOLOGY: not relevant
MOLECULE TYPE: protein
US-08-779-113-8

Query Match 61.4%; Score 35; DB 2; Length 2647;
Best Local Similarity 100.0%; Pred. NO. 1e+03; Indels 0; Gaps 0;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 YTAPOP 10
1712 YTAPOP 1717

RESULT 7
US-08-956-242-2
Sequence 2, Application US/08956242C
Patent No. 5986081

GENERAL INFORMATION:
APPLICANT: Ganetzky, Barry S.
APPLICANT: Titus, Steven A.
TITLE OF INVENTION: Polynucleotides Encoding Herg-3
FILE REFERENCE: 960296, 94550
CURRENT APPLICATION NUMBER: US/08/956,242C
CURRENT FILING DATE: 1997-10-22
NUMBER OF SEQ ID NOS: 13
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 2
LENGTH: 626
TYPE: PRT
ORGANISM: Homo sapien
FEATURE:
NAME/KEY: unsure
LOCATION: (441)..(542)
OTHER INFORMATION: Unidentified at time of filing
US-08-956-242-2

Query Match 59.6%; Score 34; DB 2; Length 626;
Best Local Similarity 60.0%; Pred. NO. 3.2e+02; Indels 0; Gaps 0;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 2 ASSTYAPQ 11
67 SSHYTFPRQ 76

RESULT 8
US-09-351-215-2
Sequence 2, Application US/09351215
Patent No. 6087488

GENERAL INFORMATION:
APPLICANT: Ganetzky, Barry S.
APPLICANT: Titus, Steven A.
TITLE OF INVENTION: Polynucleotides Encoding Herg-3
FILE REFERENCE: 960296, 94550
CURRENT APPLICATION NUMBER: US/09/351,215
CURRENT FILING DATE: 1999-07-12
EARLIER APPLICATION NUMBER: 08/956,242
EARLIER FILING DATE: 1997-10-22
NUMBER OF SEQ ID NOS: 13
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 2
LENGTH: 626
TYPE: PRT
ORGANISM: Homo sapien
FEATURE:
NAME/KEY: unsure
LOCATION: (441)..(542)
OTHER INFORMATION: Unidentified at time of filing
US-09-351-215-2

Query Match 59.6%; Score 34; DB 3; Length 626;
Best Local Similarity 60.0%; Pred. NO. 3.2e+02; Indels 0; Gaps 0;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 2 ASSTYAPQ 11
67 SSHYTFPRQ 76

RESULT 9
US-08-910-925-4
Sequence 4, Application US/08910925
Patent No. 6162601

GENERAL INFORMATION:
APPLICANT: Bandman, Olga
APPLICANT: Lal, Preeti
APPLICANT: Shah, Purvi
TITLE OF INVENTION: HUMAN PININ SPLICE VARIANT
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:

LENGTH: 209
TYPE: PRT
ORGANISM: Homo sapiens
US-09-109-100-11

Query Match
Best Local Similarity 60.0%; Score 33; DB 4; Length 209;
Pred. No. 1.5e+02;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 QASSTAPOP 10
:|::|||||
Db 149 EATAPAPOP 158

RESULT 14
US-09-109-100-12
Sequence 12, Application US/09109100C
Patent No. 6291661
GENERAL INFORMATION:
APPLICANT: Graddis, Thomas J.
APPLICANT: McGrew, Jeffrey T.
TITLE OF INVENTION: FLT3-L MUTANTS AND METHODS OF USE
FILE REFERENCE: 03260.0028
CURRENT APPLICATION NUMBER: US/09/109,100C
CURRENT FILING DATE: 1998-07-02
NUMBER OF SEQ ID NOS: 20
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 12
LENGTH: 209
TYPE: PRT
ORGANISM: Homo sapiens
US-09-109-100-12

Query Match
Best Local Similarity 60.0%; Score 33; DB 4; Length 209;
Pred. No. 1.5e+02;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 QASSTAPOP 10
:|::|||||
Db 149 EATAPAPOP 158

RESULT 15
US-09-109-100-13
Sequence 13, Application US/09109100C
Patent No. 6291661
GENERAL INFORMATION:
APPLICANT: Graddis, Thomas J.
APPLICANT: McGrew, Jeffrey T.
TITLE OF INVENTION: FLT3-L MUTANTS AND METHODS OF USE
FILE REFERENCE: 03260.0028
CURRENT APPLICATION NUMBER: US/09/109,100C
CURRENT FILING DATE: 1998-07-02
NUMBER OF SEQ ID NOS: 20
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 13
LENGTH: 209
TYPE: PRT
ORGANISM: Homo sapiens
US-09-109-100-13

Query Match
Best Local Similarity 60.0%; Score 33; DB 4; Length 209;
Pred. No. 1.5e+02;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 QASSTAPOP 10
:|::|||||
Db 149 EATAPAPOP 158

Search completed: January 13, 2003, 09:56:24
Job time : 11.5897 secs

Mon Jan 13 10:35:13 2003

GenCorp version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

January 13, 2003, 09:54:16 ; Search time 51.0513 Seconds
(without alignments)
4.180 Million cell updates/sec

Run on: US-09-554-941-2
Title: BLOSUM62
Perfect score: 1 QASSTAPPOP 11
Sequence: Gapop 10.0, Gapext 0.5

Scoring table: Gapop 10.0, Gapext 0.5

Searched: 118974 seqs, 19401057 residues

Total number of hits satisfying chosen parameters: 118974

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: Published Applications_AA:*

1: /cgn2_6/ptodata/1/pubppa/US08_NEW_PUB_PEP:*

pred. No. is the number of results predicted by chance to have a
score greater than or equal to the total score distribution.
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	52	91.2	639	10	US-09-782-906-2
2	52	91.2	639	10	US-09-782-906-3
3	52	91.2	639	10	US-09-782-906-4
4	52	91.2	639	10	US-09-782-906-5
5	38	66.7	223	10	US-09-864-761-33417
6	37	64.9	223	10	US-09-864-761-33417
7	37	64.9	223	10	US-09-864-761-33417
8	37	64.9	223	10	US-09-864-761-33417
9	36	63.2	1802	10	US-09-864-761-33691
10	35	61.4	301	9	US-09-782-906-2
11	35	61.4	301	9	US-09-782-906-3
12	34	59.6	294	10	US-09-782-906-4
13	34	59.6	294	10	US-09-782-906-5
14	34	59.6	294	10	US-09-782-906-6
15	33	57.9	209	10	US-09-904-536-9
16	33	57.9	209	10	US-09-904-536-11
17	33	57.9	209	10	US-09-904-536-12
18	33	57.9	209	10	US-09-904-536-12
19	33	57.9	209	10	US-09-904-536-12

us-09-554-941-2.rapp

20	33	57.9	209	10	US-09-904-536-13	Sequence 13, Appl
21	33	57.9	209	10	US-09-904-536-14	Sequence 14, Appl
22	33	57.9	209	10	US-09-904-536-15	Sequence 15, Appl
23	33	57.9	209	10	US-09-904-536-16	Sequence 16, Appl
24	33	57.9	209	10	US-09-904-536-17	Sequence 17, Appl
25	33	57.9	209	10	US-09-904-536-18	Sequence 18, Appl
26	33	57.9	209	10	US-09-904-536-19	Sequence 19, Appl
27	33	57.9	209	10	US-09-904-536-20	Sequence 20, Appl
28	33	57.9	209	10	US-09-904-536-21	Sequence 21, Appl
29	33	57.9	209	10	US-09-904-536-22	Sequence 22, Appl
30	33	57.9	209	10	US-09-904-536-23	Sequence 23, Appl
31	33	57.9	209	10	US-09-904-536-24	Sequence 24, Appl
32	33	57.9	209	10	US-09-904-536-25	Sequence 25, Appl
33	33	57.9	209	10	US-09-904-536-26	Sequence 26, Appl
34	33	57.9	209	10	US-09-904-536-27	Sequence 27, Appl
35	33	57.9	209	10	US-09-904-536-28	Sequence 28, Appl
36	33	57.9	209	10	US-09-904-536-29	Sequence 29, Appl
37	33	57.9	209	10	US-09-904-536-30	Sequence 30, Appl
38	33	57.9	209	10	US-09-904-536-31	Sequence 31, Appl
39	33	57.9	209	10	US-09-904-536-32	Sequence 32, Appl
40	33	57.9	209	10	US-09-904-536-33	Sequence 33, Appl
41	33	57.9	209	10	US-09-904-536-34	Sequence 34, Appl
42	33	57.9	209	10	US-09-904-536-35	Sequence 35, Appl
43	33	57.9	209	10	US-09-904-536-36	Sequence 36, Appl
44	33	57.9	209	10	US-09-904-536-37	Sequence 37, Appl
45	33	57.9	209	10	US-09-904-536-38	Sequence 38, Appl

ALIGNMENTS

RESULT 1
US-09-782-906-2
Sequence 2, Application US/09782906
Patent No. US20010051369A1
GENERAL INFORMATION:
APPLICANT: Delagrange, Simon
APPLICANT: Rittenhouse Pruss, Jennifer L.
APPLICANT: Murphy, Dennis J.
APPLICANT: Maffia III, Anthony M.
APPLICANT: Coleman, William Galactose Oxidase, Nucleic Acid Encoding Same, And Met
TITLE OF INVENTION: Variant Same
FILE REFERENCE: HER-0040
CURRENT FILING DATE: 2001-02-14
CURRENT APPLICATION NUMBER: 60/185,001
PRIOR FILING DATE: 2000-02-25
NUMBER OF SEQ ID NOS: 8
SOFTWARE: PatentIn version 3.0
SEQ ID NO 2
LENGTH: 639
ORGANISM: Dactylium dendroideus
US-09-782-906-2
Query Match
Best Local Similarity 91.2%; Score 52; DB 10; Length 639;
Matches 100.0%; Pred No. 0.14; Indels 0; Gaps 0;
Matches 0; Mismatches 0;
DB 147 QASSTAPPOP 156
RESULT 2
US-09-782-906-3
Sequence 3, Application US/09782906
Patent No. US20010051369A1
GENERAL INFORMATION:
APPLICANT: Delagrange, Simon
APPLICANT: Rittenhouse Pruss, Jennifer L.

Mon Jan 13 10:35:13 2003

us-09-554-941-2.rabp

Page 2

```

; APPLICANT: Murphy, Dennis J.
; APPLICANT: Maffia III, Anthony M.
; APPLICANT: Bylina, Edward J.
; APPLICANT: Coleman, William J.
; TITLE OF INVENTION: Variant Galactose Oxidase, Nucleic Acid Encoding Same, And Method
; FILE REFERENCE: HRP-0040
; CURRENT APPLICATION NUMBER: US/09/782,906
; PRIOR FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: 60/185,001
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3
; LENGTH: 639
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: No. US20010051369A1el Sequence
US-09-782-906-3

```

```

Query Match
Best Local Similarity 91.28; Score 52; DB 10; Length 639;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 QASSTYAPOP 10
Db 147 QASSTYAPOP 156

```

```

RESULT 3
US-09-782-906-4
; Sequence 4, Application US/09782906
; Patent No. US20010051369A1
; GENERAL INFORMATION:
; APPLICANT: Delagrave, Simon
; APPLICANT: Rittenhouse Pruss, Jennifer L.
; APPLICANT: Murphy, Dennis J.
; APPLICANT: Maffia III, Anthony M.
; APPLICANT: Bylina, Edward J.
; APPLICANT: Coleman, William J.
; TITLE OF INVENTION: Variant Galactose Oxidase, Nucleic Acid Encoding Same, And Method
; FILE REFERENCE: HRP-0040
; CURRENT APPLICATION NUMBER: US/09/782,906
; PRIOR FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: 60/185,001
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 639
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: No. US20010051369A1el Sequence
US-09-782-906-4

```

```

Query Match
Best Local Similarity 91.28; Score 52; DB 10; Length 639;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 QASSTYAPOP 10
Db 147 QASSTYAPOP 156

```

```

RESULT 4
US-09-782-906-5
; Sequence 5, Application US/09782906
; Patent No. US20010051369A1
; GENERAL INFORMATION:
; APPLICANT: Delagrave, Simon

```

```

; APPLICANT: Rittenhouse Pruss, Jennifer L.
; APPLICANT: Murphy, Dennis J.
; APPLICANT: Maffia III, Anthony M.
; APPLICANT: Bylina, Edward J.
; APPLICANT: Coleman, William J.
; TITLE OF INVENTION: Variant Galactose Oxidase, Nucleic Acid Encoding Same, And Met
; FILE REFERENCE: HRP-0040
; CURRENT APPLICATION NUMBER: US/09/782,906
; PRIOR FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: 60/185,001
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 5
; LENGTH: 639
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: No. US20010051369A1el Sequence
US-09-782-906-5

```

```

Query Match
Best Local Similarity 91.28; Score 52; DB 10; Length 639;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 QASSTYAPOP 10
Db 147 QASSTYAPOP 156

```

```

RESULT 5
US-09-864-761-33417
; Sequence 33417, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; PRIOR FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263,6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661

```

QY 1 QASSYTAPO 10
11::11111
Db 72 QATTLTAPO 81

RESULT 6
US-09-893-737-324
Sequence 324, Application US/09893737
US-09-893-737-324
Sequence 324, Application US/09893737

```

GENERAL: Sheppard, Scott R.
APPLICANT: Presnell, Scott R.
TITLE OF INVENTION: MAMMALIAN SECRETED PROTEINS
FILE REFERENCE: 00-41
CURRENT APPLICATION NUMBER: US/09/893,737
CURRENT FILING DATE: 2001-06-28
CURRENT APPLICATION NUMBER: US 60/215,446
PRIORITY FILING DATE: 2000-06-30
PRIORITY OF SEQ ID NOS: 329
NUMBER OF SEQ ID NOS: 329
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO: 324
LENGTH: 295
TYPE: prt
ORIGIN: Homo sapiens
ORIGIN: 1
ORIGIN: 2
ORIGIN: 3
ORIGIN: 4
ORIGIN: 5
ORIGIN: 6
ORIGIN: 7
ORIGIN: 8
ORIGIN: 9
ORIGIN: 10
ORIGIN: 11
ORIGIN: 12
ORIGIN: 13
ORIGIN: 14
ORIGIN: 15
ORIGIN: 16
ORIGIN: 17
ORIGIN: 18
ORIGIN: 19
ORIGIN: 20
ORIGIN: 21
ORIGIN: 22
ORIGIN: 23
ORIGIN: 24
ORIGIN: 25
ORIGIN: 26
ORIGIN: 27
ORIGIN: 28
ORIGIN: 29
ORIGIN: 30
ORIGIN: 31
ORIGIN: 32
ORIGIN: 33
ORIGIN: 34
ORIGIN: 35
ORIGIN: 36
ORIGIN: 37
ORIGIN: 38
ORIGIN: 39
ORIGIN: 40
ORIGIN: 41
ORIGIN: 42
ORIGIN: 43
ORIGIN: 44
ORIGIN: 45
ORIGIN: 46
ORIGIN: 47
ORIGIN: 48
ORIGIN: 49
ORIGIN: 50
ORIGIN: 51
ORIGIN: 52
ORIGIN: 53
ORIGIN: 54
ORIGIN: 55
ORIGIN: 56
ORIGIN: 57
ORIGIN: 58
ORIGIN: 59
ORIGIN: 60
ORIGIN: 61
ORIGIN: 62
ORIGIN: 63
ORIGIN: 64
ORIGIN: 65
ORIGIN: 66
ORIGIN: 67
ORIGIN: 68
ORIGIN: 69
ORIGIN: 70
ORIGIN: 71
ORIGIN: 72
ORIGIN: 73
ORIGIN: 74
ORIGIN: 75
ORIGIN: 76
ORIGIN: 77
ORIGIN: 78
ORIGIN: 79
ORIGIN: 80
ORIGIN: 81
ORIGIN: 82
ORIGIN: 83
ORIGIN: 84
ORIGIN: 85
ORIGIN: 86
ORIGIN: 87
ORIGIN: 88
ORIGIN: 89
ORIGIN: 90
ORIGIN: 91
ORIGIN: 92
ORIGIN: 93
ORIGIN: 94
ORIGIN: 95
ORIGIN: 96
ORIGIN: 97
ORIGIN: 98
ORIGIN: 99
ORIGIN: 100
ORIGIN: 101
ORIGIN: 102
ORIGIN: 103
ORIGIN: 104
ORIGIN: 105
ORIGIN: 106
ORIGIN: 107
ORIGIN: 108
ORIGIN: 109
ORIGIN: 110
ORIGIN: 111
ORIGIN: 112
ORIGIN: 113
ORIGIN: 114
ORIGIN: 115
ORIGIN: 116
ORIGIN: 117
ORIGIN: 118
ORIGIN: 119
ORIGIN: 120
ORIGIN: 121
ORIGIN: 122
ORIGIN: 123
ORIGIN: 124
ORIGIN: 125
ORIGIN: 126
ORIGIN: 127
ORIGIN: 128
ORIGIN: 129
ORIGIN: 130
ORIGIN: 131
ORIGIN: 132
ORIGIN: 133
ORIGIN: 134
ORIGIN: 135
ORIGIN: 136
ORIGIN: 137
ORIGIN: 138
ORIGIN: 139
ORIGIN: 140
ORIGIN: 141
ORIGIN: 142
ORIGIN: 143
ORIGIN: 144
ORIGIN: 145
ORIGIN: 146
ORIGIN: 147
ORIGIN: 148
ORIGIN: 149
ORIGIN: 150
ORIGIN: 151
ORIGIN: 152
ORIGIN: 153
ORIGIN: 154
ORIGIN: 155
ORIGIN: 156
ORIGIN: 157
ORIGIN: 158
ORIGIN: 159
ORIGIN: 160
ORIGIN: 161
ORIGIN: 162
ORIGIN: 163
ORIGIN: 164
ORIGIN: 165
ORIGIN: 166
ORIGIN: 167
ORIGIN: 168
ORIGIN: 169
ORIGIN: 170
ORIGIN: 171
ORIGIN: 172
ORIGIN: 173
ORIGIN: 174
ORIGIN: 175
ORIGIN: 176
ORIGIN: 177
ORIGIN: 178
ORIGIN: 179
ORIGIN: 180
ORIGIN: 181
ORIGIN: 182
ORIGIN: 183
ORIGIN: 184
ORIGIN: 185
ORIGIN: 186
ORIGIN: 187
ORIGIN: 188
ORIGIN: 189
ORIGIN: 190
ORIGIN: 191
ORIGIN: 192
ORIGIN: 193
ORIGIN: 194
ORIGIN: 195
ORIGIN: 196
ORIGIN: 197
ORIGIN: 198
ORIGIN: 199
ORIGIN: 200
ORIGIN: 201
ORIGIN: 202
ORIGIN: 203
ORIGIN: 204
ORIGIN: 205
ORIGIN: 206
ORIGIN: 207
ORIGIN: 208
ORIGIN: 209
ORIGIN: 210
ORIGIN: 211
ORIGIN: 212
ORIGIN: 213
ORIGIN: 214
ORIGIN: 215
ORIGIN: 216
ORIGIN: 217
ORIGIN: 218
ORIGIN: 219
ORIGIN: 220
ORIGIN: 221
ORIGIN: 222
ORIGIN: 223
ORIGIN: 224
ORIGIN: 225
ORIGIN: 226
ORIGIN: 227
ORIGIN: 228
ORIGIN: 229
ORIGIN: 230
ORIGIN: 231
ORIGIN: 232
ORIGIN: 233
ORIGIN: 234
ORIGIN: 235
ORIGIN: 236
ORIGIN: 237
ORIGIN: 238
ORIGIN: 239
ORIGIN: 240
ORIGIN: 241
ORIGIN: 242
ORIGIN: 243
ORIGIN: 244
ORIGIN: 245
ORIGIN: 246
ORIGIN: 247
ORIGIN: 248
ORIGIN: 249
ORIGIN: 250
ORIGIN: 251
ORIGIN: 252
ORIGIN: 253
ORIGIN: 254
ORIGIN: 255
ORIGIN: 256
ORIGIN: 257
ORIGIN: 258
ORIGIN: 259
ORIGIN: 260
ORIGIN: 261
ORIGIN: 262
ORIGIN: 263
ORIGIN: 264
ORIGIN: 265
ORIGIN: 266
ORIGIN: 267
ORIGIN: 268
ORIGIN: 269
ORIGIN: 270
ORIGIN: 271
ORIGIN: 272
ORIGIN: 273
ORIGIN: 274
ORIGIN: 275
ORIGIN: 276
ORIGIN: 277
ORIGIN: 278
ORIGIN: 279
ORIGIN: 280
ORIGIN: 281
ORIGIN: 282
ORIGIN: 283
ORIGIN: 284
ORIGIN: 285
ORIGIN: 286
ORIGIN: 287
ORIGIN: 288
ORIGIN: 289
ORIGIN: 290
ORIGIN: 291
ORIGIN: 292
ORIGIN: 293
ORIGIN: 294
ORIGIN: 295
ORIGIN: 296
ORIGIN: 297
ORIGIN: 298
ORIGIN: 299
ORIGIN: 300
ORIGIN: 301
ORIGIN: 302
ORIGIN: 303
ORIGIN: 304
ORIGIN: 305
ORIGIN: 306
ORIGIN: 307
ORIGIN: 308
ORIGIN: 309
ORIGIN: 310
ORIGIN: 311
ORIGIN: 312
ORIGIN: 313
ORIGIN: 314
ORIGIN: 315
ORIGIN: 316
ORIGIN: 317
ORIGIN: 318
ORIGIN: 319
ORIGIN: 320
ORIGIN: 321
ORIGIN: 322
ORIGIN: 323
ORIGIN: 324
ORIGIN: 325
ORIGIN: 326
ORIGIN: 327
ORIGIN: 328
ORIGIN: 329
ORIGIN: 330
ORIGIN: 331
ORIGIN: 332
ORIGIN: 333
ORIGIN: 334
ORIGIN: 335
ORIGIN: 336
ORIGIN: 337
ORIGIN: 338
ORIGIN: 339
ORIGIN: 340
ORIGIN: 341
ORIGIN: 342
ORIGIN: 343
ORIGIN: 344
ORIGIN: 345
ORIGIN: 346
ORIGIN: 347
ORIGIN: 348
ORIGIN: 349
ORIGIN: 350
ORIGIN: 351
ORIGIN: 352
ORIGIN: 353
ORIGIN: 354
ORIGIN: 355
ORIGIN: 356
ORIGIN: 357
ORIGIN: 358
ORIGIN: 359
ORIGIN: 360
ORIGIN: 361
ORIGIN: 362
ORIGIN: 363
ORIGIN: 364
ORIGIN: 365
ORIGIN: 366
ORIGIN: 367
ORIGIN: 368
ORIGIN: 369
ORIGIN: 370
ORIGIN: 371
ORIGIN: 372
ORIGIN: 373
ORIGIN: 374
ORIGIN: 375
ORIGIN: 376
ORIGIN: 377
ORIGIN: 378
ORIGIN: 379
ORIGIN: 380
ORIGIN: 381
ORIGIN: 382
ORIGIN: 383
ORIGIN: 384
ORIGIN: 385
ORIGIN: 386
ORIGIN: 387
ORIGIN: 388
ORIGIN: 389
ORIGIN: 390
ORIGIN: 391
ORIGIN: 392
ORIGIN: 393
ORIGIN: 394
ORIGIN: 395
ORIGIN: 396
ORIGIN: 397
ORIGIN: 398
ORIGIN: 399
ORIGIN: 400
ORIGIN: 401
ORIGIN: 402
ORIGIN: 403
ORIGIN: 404
ORIGIN: 405
ORIGIN: 406
ORIGIN: 407
ORIGIN: 408
ORIGIN: 409
ORIGIN: 410
ORIGIN: 411
ORIGIN: 412
ORIGIN: 413
ORIGIN: 414
ORIGIN: 415
ORIGIN: 416
ORIGIN: 417
ORIGIN: 418
ORIGIN: 419
ORIGIN: 420
ORIGIN: 421
ORIGIN: 422
ORIGIN: 423
ORIGIN: 424
ORIGIN: 425
ORIGIN: 426
ORIGIN: 427
ORIGIN: 428
ORIGIN: 429
ORIGIN: 430
ORIGIN: 431
ORIGIN: 432
ORIGIN: 433
ORIGIN: 434
ORIGIN: 435
ORIGIN: 436
ORIGIN: 437
ORIGIN: 438
ORIGIN: 439
ORIGIN: 440
ORIGIN: 441
ORIGIN: 442
ORIGIN: 443
ORIGIN: 444
ORIGIN: 445
ORIGIN: 446
ORIGIN: 447
ORIGIN: 448
ORIGIN: 449
ORIGIN: 450
ORIGIN: 451
ORIGIN: 452
ORIGIN: 453
ORIGIN: 454
ORIGIN: 455
ORIGIN: 456
ORIGIN: 457
ORIGIN: 458
ORIGIN: 459
ORIGIN: 460
ORIGIN: 461
ORIGIN: 462
ORIGIN: 463
ORIGIN: 464
ORIGIN: 465
ORIGIN: 466
ORIGIN: 467
ORIGIN: 468
ORIGIN: 469
ORIGIN: 470
ORIGIN: 471
ORIGIN: 472
ORIGIN: 473
ORIGIN: 474
ORIGIN: 475
ORIGIN: 476
ORIGIN: 477
ORIGIN: 478
ORIGIN: 479
ORIGIN: 480
ORIGIN: 481
ORIGIN: 482
ORIGIN: 483
ORIGIN: 484
ORIGIN: 485
ORIGIN: 486
ORIGIN: 487
ORIGIN: 488
ORIGIN: 489
ORIGIN: 490
ORIGIN: 491
ORIGIN: 492
ORIGIN: 493
ORIGIN: 494
ORIGIN: 495
ORIGIN: 496
ORIGIN: 497
ORIGIN: 498
ORIGIN: 499
ORIGIN: 500
ORIGIN: 501
ORIGIN: 502
ORIGIN: 503
```

Query Match	Similarity	85.78;	Percent	0;
Best Local	Similarity	1;	Mismatches	0;
Matches	6;	Conservative		

QY 5 YTAPOPO 11
111111:
Db 187 YTAPOPE 193

RESULT 7
US-09-729-674-160
US-09-729-674-160 Application US/09729674
; Sequence 160, Application US3335A1
; Patent No. US20010039335A1
; GENERAL INFORMATION:
; GENERAL INFORMATION:

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

QY 5 YTAPOPO 11
 89 YTAPOPE 95
 Db

RESULT 8
US-09-965-553-8
Application US/09965553
sequence 8, accession 12259A1

Patent NO. INFORMATION:
GENERAL INFORMATION: David A.
APPLICANT: Wright, Daniel F.
APPLICANT: Voytas, Daniel
TITLE OF INVENTION: Plant Retroelements and Methods Related Thereto
FILE REFERENCE: P-1065 ISRP Plant Retroelement
CURRENT FILING DATE: 2001-09-27
PRIOR APPLICATION NUMBER: 09/322,478
PRIOR FILING DATE: 1999-05-28
PRIOR APPLICATION NUMBER: 60/087125
PRIOR FILING DATE: 1998-05-29
NUMBER OF SEQ ID NOS: 41
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO: 8
LENGTH: 532
TYPE: prt
ORGANISM: Artificial Sequence
FEATURE: Description of Artificial Sequence: plant
OTHER INFORMATION: retroelement sequence
OS-09-965-553-8

63.2%; Score 36; DB 10; Length 532;
GC 60; Indels 0; Gaps 0;

```

Query Match Similarity 60.0%;
Best Local Similarity 2;
Matches 6; Conservative

```

```

QY      1 QASSYTAPOP 10
        : ||||:| |
Db      36 EGSSYTSPPP 45

```

RESULT 9

Mon Jan 13 10:35:13 2003

us-09-554-941-2.rapb

Page 4

```
US-09-965-553-18
; Sequence 18, Application US/09965553
; Patent No. US20020112259A1
; GENERAL INFORMATION:
; APPLICANT: Wright, David A.
; TITLE OF INVENTION: Plant Retroelements and Methods Related Thereto
; FILE REFERENCE: P-1065 ISUP Plant Retroelement
; CURRENT FILING DATE: 2001-09-27
; PRIOR FILING DATE: 1999-05-28/322,478
; PRIOR APPLICATION NUMBER: 09/322,478
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: Patent Ver. 2.0
; SEQ ID NO: 18
; LENGTH: 1802
; TYPE: PRT
; ORGANISM: Artificial Sequence
; OTHER INFORMATION: Description of Artificial Sequence: Plant
US-09-965-553-18
Query Match
Best Local Similarity 63.2%; Score 36; DB 10; Length 1802;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
QY 1 QASSTAPOP 10
DB 36 EGSSYTSPP 45
RESULT 10
US-09-864-761-35691
; Sequence 35691, Application US/09864761
; Patent No. US2002004675A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Neomulca-X-1
; CURRENT FILING DATE: 2001-05-23
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: CB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
```

```
US-09-738-626-6745
; Sequence 6745, Application US/09738626
; Patent No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIRO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, MOKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT FILING DATE: 2000-12-18
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: JP 00/280988
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: Patent Ver. 3.0
; SEQ ID NO: 6745
; LENGTH: 301
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-6745
Query Match
Best Local Similarity 77.8%; Score 35; DB 10; Length 53;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 2 ASSYTAPOP 10
DB 44 ASSYTAPOP 52
RESULT 11
US-09-738-626-6745
; Sequence 6745, Application US/09738626
; Patent No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIRO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, MOKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT FILING DATE: 2000-12-18
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: JP 00/280988
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: Patent Ver. 3.0
; SEQ ID NO: 6745
; LENGTH: 301
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
```

Query Match 61.4%; Score 35; DB 9; Length 301;
Best Local Similarity 87.5%; Pred. No. 49;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 ASSYTAPO 9
11111111
Db 38 ASSYTMPQ 45

RESULT 12

US-09-764-864-1515
; Sequence 1515, Application US/09764864
; Patent No. US20020132753A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT223
; CURRENT APPLICATION NUMBER: US/09/764,864
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1792
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 1515
; LENGTH: 294
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (84)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-864-1515

Query Match 59.6%; Score 34; DB 10; Length 294;
Best Local Similarity 54.5%; Pred. No. 71;
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 QASSTAPQPO 11
11111111
Db 181 QSSSPTSPEPR 191

RESULT 13

US-09-764-864-947
; Sequence 947, Application US/09764864
; Patent No. US20020132753A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT223
; CURRENT APPLICATION NUMBER: US/09/764,864
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1792
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 947
; LENGTH: 323
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-764-864-947

Query Match 59.6%; Score 34; DB 10; Length 323;
Best Local Similarity 54.5%; Pred. No. 78;
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 QASSTAPQPO 11
11111111
Db 210 QSSSPTSPEPR 220

RESULT 14

US-09-764-864-1095
; Sequence 1095, Application US/09764864
; Patent No. US20020132753A1

GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT223
; CURRENT APPLICATION NUMBER: US/09/764,864
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1792
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 1095
; LENGTH: 323
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-764-864-1095

Query Match 59.6%; Score 34; DB 10; Length 323;
Best Local Similarity 54.5%; Pred. No. 78;
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 QASSTAPQPO 11
11111111
Db 210 QSSSPTSPEPR 220

RESULT 15

US-09-738-626-5433
; Sequence 5433, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAOKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: Patentln ver. 3.0
; SEQ ID NO 5433
; LENGTH: 407
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-5433

Query Match 59.6%; Score 34; DB 9; Length 407;
Best Local Similarity 85.7%; Pred. No. 99;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 YRAPQPO 11
11111111
Db 254 YQAPQPO 260

Search completed: January 13, 2003, 10:02:49
Job time: 51.0513 secs

Mon Jan 13 10:35:13 2003

us-09-554-941-2.rag

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 13, 2003, 09:53:15 (Search time 25.1026 Seconds
(without alignments)
58.391 Million cell updates/sec)

Title: US-09-554-941-2
Perfect score: 57
Sequence: 1 QASSTAPQ 11

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 908470 seqs, 133250620 residues 908470

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: 1: A_Geneseq_101002:*

2: /SID2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
3: /SID2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
4: /SID2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
5: /SID2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
6: /SID2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
7: /SID2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
8: /SID2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
9: /SID2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10: /SID2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
11: /SID2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12: /SID2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13: /SID2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14: /SID2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15: /SID2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16: /SID2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17: /SID2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18: /SID2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19: /SID2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20: /SID2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
21: /SID2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22: /SID2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
23: /SID2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	57	100.0	11	AA122919	Galactose oxidase
2	52	91.2	639	AAE12925	Dactylium dendroid
3	52	91.2	639	AAE12926	Dactylium dendroid
4	52	91.2	639	AAE12927	Dactylium dendroid
5	52	91.2	639	AAE12928	Dactylium dendroid
6	52	91.2	639	AAE12929	Dactylium dendroid
7	52	91.2	639	AAE12930	Dactylium dendroid
8	52	91.2	639	AAE12931	Dactylium dendroid
9	52	91.2	639	AAE12932	Dactylium dendroid
10	52	91.2	639	AAE12933	Dactylium dendroid

11	52	91.2	639	22	AAE12934	Dactylium dendroid
12	52	91.2	639	22	AAE12935	Dactylium dendroid
13	52	91.2	639	22	AAE12936	Dactylium dendroid
14	52	91.2	639	22	AAE12937	Dactylium dendroid
15	52	91.2	639	22	AAE12938	Dactylium dendroid
16	52	91.2	639	22	AAE12939	Dactylium dendroid
17	52	91.2	639	22	AAE12940	Dactylium dendroid
18	52	91.2	639	22	AAE12941	Dactylium dendroid
19	52	91.2	639	22	AAE12942	Dactylium dendroid
20	52	91.2	639	22	AAE12943	Dactylium dendroid
21	52	91.2	639	22	AAE12944	Dactylium dendroid
22	52	91.2	639	22	AAE12945	Dactylium dendroid
23	52	91.2	639	22	AAE12946	Dactylium dendroid
24	52	91.2	639	22	AAE12947	Dactylium dendroid
25	52	91.2	639	22	AAE12948	Dactylium dendroid
26	52	91.2	639	22	AAE12949	Dactylium dendroid
27	52	91.2	639	22	AAE12950	Dactylium dendroid
28	52	91.2	639	22	AAE12951	Dactylium dendroid
29	52	91.2	639	22	AAE12952	Dactylium dendroid
30	52	91.2	639	22	AAE12953	Dactylium dendroid
31	52	91.2	639	22	AAE12954	Dactylium dendroid
32	52	91.2	639	22	AAE12955	Dactylium dendroid
33	52	91.2	639	22	AAE12956	Dactylium dendroid
34	52	91.2	639	22	AAE12957	Dactylium dendroid
35	52	91.2	639	22	AAE12958	Dactylium dendroid
36	52	91.2	639	22	AAE12959	Dactylium dendroid
37	52	91.2	639	22	AAE12960	Dactylium dendroid
38	52	91.2	639	22	AAE12961	Dactylium dendroid
39	52	91.2	639	22	AAE12962	Dactylium dendroid
40	52	91.2	639	22	AAE12963	Dactylium dendroid
41	52	91.2	639	22	AAE12964	Dactylium dendroid
42	52	91.2	639	22	AAE12965	Dactylium dendroid
43	52	91.2	639	22	AAE12966	Dactylium dendroid
44	52	91.2	639	22	AAE12967	Dactylium dendroid
45	52	91.2	639	22	AAE12968	Dactylium dendroid

ALIGNMENTS

RESULT 1
AA122919 standard; peptide; 11 AA.

AA122919: 19-AUG-1999 (first entry)

Galactose oxidase linker used to make proteinase inhibitor fusions.

Peptide linker: fusion protein; pathogen resistance; pathogen tolerance; plant; transgene; proteinase inhibitor; Arabidopsis thaliana; parasitic nematode.

Synthetic.

W09928484-A1.

10-JUN-1999.

01-DEC-1998; 98WO-EP07792.

03-DEC-1997; 97GB-0025556.

(NOVS) NOVARTIS AG.

Atkinson HJ, McPherson MJ, Urwin PE;

WPI; 1999-385387/32.

Proteinase inhibitor fusion proteins

Claim 9; Page 23; 39pp; English.

XX The present sequence represents a peptide linker used in the fusion
 CC proteins of the invention. The specification describes a method for
 CC improving pathogen resistance or tolerance of a plant. The method
 CC comprises transformation of the plant with a transgene encoding a
 CC fusion protein which comprises two or more protein or domains that
 CC are capable of improving pathogen resistance or tolerance when
 CC expressed on their own. Specifically, two distinct proteinase
 CC inhibitors are co-delivered, as a fusion, to Arabidopsis thaliana. The
 CC proteins or domains are connected by a peptide linker. The method is
 CC used to improve pathogen resistance or tolerance of a plant and its
 CC descendants, especially against parasitic nematode attack.

SO Sequence 11 AA;

Query Match 100.0%; Score 57; DB 20; Length 11;
 Best Local Similarity 100.0%; Pred. No. 0.0018;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 QASSTYAPQPQ 11
 Db 1 QASSTYAPQPQ 11

RESULT 2
 ID AAE12925 standard; Protein: 639 AA.
 AC AAE12925;
 XX
 DT 15-JAN-2002 (first entry)
 DE Dactylium dendroides wild type galactose oxidase (GO) mature protein.

XX Galactose oxidase; GO: paper making; aldehyde synthesis; detection;
 KW pulp biobleaching; colorectal adenocarcinoma; dental plaque.
 OS Dactylium dendroides.
 XX
 PN WO200162938-A2.
 PD 30-AUG-2001.

XX 21-FEB-2001; 2001WO-US05732.
 PF 25-FEB-2000; 2000US-185001P.
 PR 14-FEB-2001; 2001US-0782906.
 XX
 PA (HERC) HERCULES INC.
 XX

PI Maffia AM, Delagrave S, Murphy DJ, Rittenhouse Pruss J, Bylina E;
 PI Coleman WJ;
 DR WPI: 2001-648219/74.
 DR N-PSDB; AAD21273.
 XX

PT Novel variant galactose oxidase enzyme for oxidizing a compound such as
 PT guar which is useful in paper making, for enzymatic synthesis of
 PT aldehydes and pulp biobleaching, has at least one substituted amino
 acid -
 XX
 PS
 XX

Example 1; Page 10; 65pp; English.

CC The invention relates to variant galactose oxidase (vGO) proteins and
 CC their cDNA molecules. Galactose oxidases are useful for oxidizing
 CC D-galactose side chains of many complex compounds such as guar, major
 CC glycolipid of human red cells. GO is useful in paper making processes,
 CC generation of H2O2 in situ, enzymatic synthesis of aldehydes, pulp
 CC biobleaching, the use of GO-Schiff's reagent for early detection, pulp
 CC prognosis in human colorectal adenocarcinoma and use of GO-glucan
 CC binding domain fusion proteins as targeting inhibitors of dental
 CC plaque bacteria. GO is useful as a research tool for identification,
 CC characterisation and purification of interacting regulatory proteins.

CC The present sequence is Dactylium dendroides wild type galactose
 CC oxidase (GO) mature protein.
 CC
 XX
 SO Sequence 639 AA;

Query Match 91.2%; Score 52; DB 22; Length 639;
 Best Local Similarity 100.0%; Pred. No. 0.82;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 QASSTYAPQP 10
 Db 147 QASSTYAPQP 156

RESULT 3
 ID AAE12926 standard; Protein: 639 AA.
 AC AAE12926;
 XX
 DT 15-JAN-2002 (first entry)
 DE Dactylium dendroides galactose oxidase GO8-1 variant.

XX Galactose oxidase; GO: paper making; aldehyde synthesis; detection;
 KW pulp biobleaching; colorectal adenocarcinoma; dental plaque; mutant;
 XX muten; variant.
 OS Synthetic.
 XX
 PN WO200162938-A2.
 PD 30-AUG-2001.

XX 21-FEB-2001; 2001WO-US05732.
 PF 25-FEB-2000; 2000US-185001P.
 PR 14-FEB-2001; 2001US-0782906.
 XX
 PA (HERC) HERCULES INC.
 XX
 PI Maffia AM, Delagrave S, Murphy DJ, Rittenhouse Pruss J, Bylina E;
 PI Coleman WJ;
 DR WPI: 2001-648219/74.
 DR N-PSDB; AAD21274.
 XX

PT Novel variant galactose oxidase enzyme for oxidizing a compound such as
 PT guar which is useful in paper making, for enzymatic synthesis of
 PT aldehydes and pulp biobleaching, has at least one substituted amino
 acid -
 XX
 PS
 XX

Claim 10; Page 45; 65pp; English.

CC The invention relates to variant galactose oxidase (vGO) proteins and
 CC their cDNA molecules. Galactose oxidases are useful for oxidizing
 CC D-galactose side chains of many complex compounds such as guar, major
 CC glycolipid of human red cells. GO is useful in paper making processes,
 CC generation of H2O2 in situ, enzymatic synthesis of aldehydes, pulp
 CC biobleaching, the use of GO-Schiff's reagent for early detection, pulp
 CC prognosis in human colorectal adenocarcinoma and use of GO-glucan
 CC binding domain fusion proteins as targeting inhibitors of dental
 CC plaque bacteria. GO is useful as a research tool for identification,
 CC characterisation and purification of interacting regulatory proteins.

Mon Jan 13 10:35:13 2003

us-09-554-941-2.rag

Page 3

The present sequence is Dactylium dendroides galactose oxidase (GO)

CC variant.

Sequence 639 AA; 91.2%; Score 52; DB 22; Length 639; Gaps 0;

Query Match 100.0%; Pred. No. 0.82; 0; Indels 0; Gaps 0;

Best Local Similarity 100.0%; Pred. No. 0.82; 0; Indels 0; Gaps 0;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 QASSTAPOP 10

147 QASSTAPOP 156

RESULT 4

AAE12927 standard; Protein: 639 AA.

AAE12927:

15-JAN-2002 (first entry)

Dactylium dendroides galactose oxidase GOB-1H3A variant;

Galactose oxidase; GO; paper making; aldehyde synthesis; detection;

pulp biobleaching; colorectal adenocarcinoma; dental plaque; variant;

mutant; muten.

Dactylium dendroides.

Synthetic.

Location/Qualifiers

Key Misc-difference 63 /note= "Wild type Gln substituted with Lys"

Misc-difference 195 /note= "Wild type Gly substituted with Ala"

Misc-difference 383 /note= "Wild type Cys substituted with Ser"

Misc-difference 436 /note= "Wild type Tyr substituted with His"

Misc-difference 494 /note= "Wild type Val substituted with Ala"

WO200162938-A2.

30-AUG-2001.

21-FEB-2001; 2001MO-US05732.

25-FEB-2000; 2000US-185001P.

14-FEB-2001; 2001US-0782906.

(HERC) HERCULES INC.

Maffia AM, Delagrave S, Murphy DJ, Rittenhouse Pruss J, Bylina E;

Coleman WJ;

WPI: 2001-648219/74.

N-PSDB: AAD21275.

Novel variant galactose oxidase enzyme for oxidizing a compound such as

quar which is useful in paper making, has at least one substituted amino

aldehydes and pulp biobleaching, has at least one substituted amino

acid

Claim 10: Page 46; 65pp; English.

The invention relates to variant galactose oxidase (GO) proteins and

their CDNA molecules. Galactose oxidases are useful for oxidizing

D-galactose side chains of many complex compounds such as guar, major

glycolipid of human red cells. GO is useful in paper making processes,

generation of H2O2 in situ, enzymatic synthesis of aldehydes, pulp

biobleaching, the use of GO-Schiff's reagent for early detection and

prognosis in human colorectal adenocarcinoma and use of GO-glucan

binding domain fusion proteins as targeting inhibitors of dental

plaque bacteria. GO is useful as a research tool for identifying

characterisation and purification of interacting regulatory proteins.

The present sequence is Dactylium dendroides galactose oxidase (GO)

variant.

Sequence 639 AA;

91.2%; Score 52; DB 22; Length 639; Gaps 0;

Query Match 100.0%; Pred. No. 0.82; 0; Indels 0; Gaps 0;

Best Local Similarity 100.0%; Pred. No. 0.82; 0; Indels 0; Gaps 0;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 QASSTAPOP 10

147 QASSTAPOP 156

15-JAN-2002 (first entry)

Dactylium dendroides galactose oxidase (GO) 7.3.2 variant.

Galactose oxidase; GO; paper making; aldehyde synthesis; detection;

pulp biobleaching; colorectal adenocarcinoma; dental plaque; variant;

mutant; muten.

Dactylium dendroides.

Synthetic.

Location/Qualifiers

Key Misc-difference 248 /note= "Wild type Lys substituted with Glu"

Misc-difference 352 /note= "Wild type Thr substituted with Ser"

Misc-difference 366 /note= "Wild type Lys substituted with Arg"

Misc-difference 383 /note= "Wild type Cys substituted with Ser"

Misc-difference 436 /note= "Wild type Tyr substituted with His"

Misc-difference 494 /note= "Wild type Val substituted with Ala"

WO200162938-A2.

30-AUG-2001.

21-FEB-2001; 2001MO-US05732.

25-FEB-2000; 2000US-185001P.

14-FEB-2001; 2001US-0782906.

(HERC) HERCULES INC.

Maffia AM, Delagrave S, Murphy DJ, Rittenhouse Pruss J, Bylina E;

Coleman WJ;

WPI: 2001-648219/74.

N-PSDB: AAD21276.

Novel variant galactose oxidase enzyme for oxidizing a compound such as

quar which is useful in paper making, has at least one substituted amino

aldehydes and pulp biobleaching, has at least one substituted amino

acid

Claim 10: Page 46; 65pp; English.

CC The invention relates to variant galactose oxidase (VGO) proteins and
 CC their cDNA molecules. Galactose oxidases are useful for oxidizing
 CC D-galactose side chains of many complex compounds such as guar, major
 CC glycolipid of human red cells. GO is useful in paper making processes,
 CC generation of H₂O₂ in situ, enzymatic synthesis of aldehydes, pulp
 CC biobleaching, the use of GO-Schiff's reagent for early detection and
 CC prognosis in human colorectal adenocarcinoma and use of GO-glucan
 CC binding domain fusion proteins as targeting inhibitors of dental
 CC plaque bacteria. GO is useful as a research tool for identification
 CC characterization and purification of interacting regulatory proteins.
 CC The present sequence is Dactylium dendroides galactose oxidase (GO).
 CC variant.

SO Sequence 639 AA;

Query Match

Best Local Similarity 91.2%; Score 52; DB 22; Length 639;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 QASSTYAPQP 10
 DB 147 QASSTYAPQP 156

RESULT 6

AAE12929

ID AAE12929 standard; Protein; 639 AA.

AC AAE12929;

DT 15-JAN-2002 (first entry)

DE Dactylium dendroides galactose oxidase variant, C383S/Y436N.

KW Galactose oxidase; GO; paper making; aldehyde synthesis; detection;
 KW pulp biobleaching; colorectal adenocarcinoma; dental plaque; mutant;
 KW muten; variant.

OS Dactylium dendroides.
 OS Synthetic.

Key Location/Qualifiers
 FT Misc-difference 383 /note= "Wild type Cys substituted with Ser"
 FT Misc-difference 436 /note= "Wild type Tyr substituted with Asn"
 FT Misc-difference 494 /note= "Wild type Val substituted with Ala"

MO200162938-A2.

30-AUG-2001.

21-FEB-2001; 2001MO-US05732.

25-FEB-2000; 2000US-185001P.

14-FEB-2001; 2001US-0782906.

(HERC) HERCULES INC.

Maffia AM, Delagrave S, Murphy DJ, Rittenhouse Pruss J, Bylina E;
 Coleman WJ;

WPI: 2001-648219/74.

Novel variant galactose oxidase enzyme for oxidizing a compound such as
 guar which is useful in paper making, for enzymatic synthesis of
 aldehydes and pulp biobleaching, has at least one substituted amino
 acid

Claim 7; Page -: 65pp; English.

CC glycolipid of human red cells. GO is useful in paper making processes,
 CC generation of H₂O₂ in situ, enzymatic synthesis of aldehydes, pulp
 CC biobleaching, the use of GO-Schiff's reagent for early detection and
 CC prognosis in human colorectal adenocarcinoma and use of GO-glucan
 CC binding domain fusion proteins as targeting inhibitors of dental
 CC plaque bacteria. GO is useful as a research tool for identification
 CC characterization and purification of interacting regulatory proteins.
 CC The present sequence is Dactylium dendroides galactose oxidase (GO).
 CC variant.
 CC Note: The present sequence is not shown in the specification but is
 CC derived from Dactylium dendroides galactose oxidase spq ID NO: 2
 CC shown in page 10 of the specification.

SO Sequence 639 AA;

Query Match

Best Local Similarity 91.2%; Score 52; DB 22; Length 639;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 QASSTYAPQP 10
 DB 147 QASSTYAPQP 156

RESULT 7

AAE12930

ID AAE12930 standard; Protein; 639 AA.

AC AAE12930;

DT 15-JAN-2002 (first entry)

DE Dactylium dendroides galactose oxidase variant, C383S/Y436N/V494A/Q63K.

KW Galactose oxidase; GO; paper making; aldehyde synthesis; detection;
 KW pulp biobleaching; colorectal adenocarcinoma; dental plaque; mutant;
 KW muten; variant.

OS Dactylium dendroides.
 OS Synthetic.

Key Location/Qualifiers
 FT Misc-difference 63 /note= "Wild type Gln substituted with Lys"
 FT Misc-difference 383 /note= "Wild type Cys substituted with Ser"
 FT Misc-difference 436 /note= "Wild type Tyr substituted with Asn"
 FT Misc-difference 494 /note= "Wild type Val substituted with Ala"

MO200162938-A2.

30-AUG-2001.

21-FEB-2001; 2001MO-US05732.

25-FEB-2000; 2000US-185001P.

14-FEB-2001; 2001US-0782906.

(HERC) HERCULES INC.

Maffia AM, Delagrave S, Murphy DJ, Rittenhouse Pruss J, Bylina E;
 Coleman WJ;

WPI: 2001-648219/74.

Novel variant galactose oxidase enzyme for oxidizing a compound such as
 guar which is useful in paper making, for enzymatic synthesis of
 aldehydes and pulp biobleaching, has at least one substituted amino
 acid

Claim 9; Page -: 65pp; English.

PT	aldehydes and pulp biobleaching, has at least one substituted amino
PT	acid -
XX	
PS	Claim 18; Page -; 65pp; English.
XX	
CC	The invention relates to variant galactose oxidase (vGO) proteins and
CC	their cDNA molecules. Galactose oxidases are useful for oxidising
CC	D-galactose side chains of many complex compounds such as guar, major
CC	glycolipid of human red cells. GO is useful in paper making processes,
CC	generation of H2O2 in situ, enzymatic synthesis of aldehydes, pulp
CC	biobleaching, the use of GO-Schiff's reagent for early detection and
CC	prognosis in human colorectal adenocarcinoma and use of GO-glucan
CC	binding domain fusion proteins as targeting inhibitors of dental
CC	plaque bacteria. GO is useful as a research tool for identification,
CC	characterisation and purification of interacting regulatory proteins.
CC	The present sequence is Dactylium dendroides galactose oxidase (GO)
CC	variant.
CC	Note: The present sequence is not shown in the specification but is
CC	derived from Dactylium dendroides galactose oxidase SEQ ID NO: 2
CC	shown in page 10 of the specification.
XX	
SO	Sequence 639 AA;
QY	1 QASSTYAPQP 10
Db	147 QASSTYAPQP 156
RESULT 9	
AAEI2932	
ID	AAEI2932 standard; Protein; 639 AA.
XX	
NC	AAEI2932;
DT	15-JAN-2002 (first entry)
XX	
DE	Dactylium dendroides galactose oxidase variant, C383S/Y436H.
XX	
KW	Galactose oxidase; GO; paper making; aldehyde synthesis; detection;
KW	pulp biobleaching; colorectal adenocarcinoma; dental plaque; mutant;
KW	muten; variant.
XX	
OS	Dactylium dendroides.
OS	Synthetic.
XX	
FH	Key
FT	Misc-difference 383
FT	/note= "Wild type Cys substituted with Ser"
FT	Misc-difference 436
FT	/note= "Wild type Tyr substituted with His"
XX	
XX	WO200162938-A2.
XX	
XX	30-AUG-2001.
PD	
XX	
PT	21-FEB-2001; 2001WO-US05732.
XX	
XX	25-FEB-2000; 2000US-185001P.
PR	14-FEB-2001; 2001US-0782906.
XX	
XX	
PA	(HERC) HERCULES INC.
XX	
PI	Maffia AM, Delagrave S, Murphy DJ, Rittenhouse Pruss J, Bylina E;
XX	Coleman WJ;
XX	
DR	WPI; 2001-648219/74.
XX	
PT	Novel variant galactose oxidase enzyme for oxidizing a compound such as
PT	guar which is useful in paper making, for enzymatic synthesis of

PT aldehydes and pulp biobleaching, has at least one substituted amino
PT acid -
XX
PS Claim 19; Page -: 65pp; English.
XX
CC The invention relates to variant galactose oxidase (VGO) proteins and
CC their cDNA molecules. Galactose oxidases are useful for oxidizing
CC D-galactose side chains of many complex compounds such as guar, major
CC glycolipid of human red cells. GO is useful in paper making processes,
CC generation of H2O2 in situ, enzymatic synthesis of aldehydes, pulp
CC biobleaching, the use of GO-Schiff's reagent for early detection and
CC prognosis in human colorectal adenocarcinoma and use of GO-glucan
CC binding domain fusion proteins as targeting inhibitors of dental
CC plaque bacteria. GO is useful as a research tool for identification,
CC characterisation and purification of interacting regulatory proteins.
CC The present sequence is Dactylium dendroides galactose oxidase (GO)
CC variant.
CC Note: The present sequence is not shown in the specification but is
CC derived from Dactylium dendroides galactose oxidase SEQ ID NO: 2
CC shown in page 10 of the specification.
XX
SQ Sequence 639 AA;
Query Match 91.2%; Score 52; DB 22; Length 639;
Best Local Similarity 100.0%; Pred. No. 0.82;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 QASSTYAPQP 10
IIIIIIIIII
DB 147 QASSTYAPQP 156
RESULT 10
AAE12933
ID AAE12933 standard; Protein; 639 AA.
XX
AC AAE12933;
XX
DT 15-JAN-2002 (first entry)
XX
DE Dactylium dendroides galactose oxidase GO.05h1B variant.
XX
KM Galactose oxidase; GO; paper making; aldehyde synthesis; detection;
KM pulp biobleaching; colorectal adenocarcinoma; dental plaque; mutant;
KM mutein; variant.
XX
OS Dactylium dendroides.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 115 /note= "Wild type Asn substituted with His"
FT
XX
PN WO200162938-A2.
XX
PD 30-AUG-2001.
XX
PF 21-FEB-2001; 2001WO-US05732.
XX
PR 25-FEB-2000; 2000US-185001P.
PR 14-FEB-2001; 2001US-0782906.
XX
PA (HERC) HERCULES INC.
XX
PI Maffia AM, Delagrave S, Murphy DJ, Rittenhouse Pruss J, Bylina E;
PI Coleman WJ;
XX
DR WPI; 2001-648219/74.
XX
PT Novel variant galactose oxidase enzyme for oxidizing a compound such as
PT guar which is useful in paper making, for enzymatic synthesis of
PT aldehydes and pulp biobleaching, has at least one substituted amino
PT acid -

XX
PS Example 1; Page -: 65pp; English.
XX
CC The invention relates to variant galactose oxidase (VGO) proteins and
CC their cDNA molecules. Galactose oxidases are useful for oxidizing
CC D-galactose side chains of many complex compounds such as guar, major
CC glycolipid of human red cells. GO is useful in paper making processes,
CC generation of H2O2 in situ, enzymatic synthesis of aldehydes, pulp
CC biobleaching, the use of GO-Schiff's reagent for early detection and
CC prognosis in human colorectal adenocarcinoma and use of GO-glucan
CC binding domain fusion proteins as targeting inhibitors of dental
CC plaque bacteria. GO is useful as a research tool for identification,
CC characterisation and purification of interacting regulatory proteins.
CC The present sequence is Dactylium dendroides galactose oxidase (GO)
CC variant.
CC Note: The present sequence is not shown in the specification but is
CC derived from Dactylium dendroides galactose oxidase SEQ ID NO: 2
CC shown in page 10 of the specification.
XX
SQ Sequence 639 AA;
Query Match 91.2%; Score 52; DB 22; Length 639;
Best Local Similarity 100.0%; Pred. No. 0.82;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 QASSTYAPQP 10
IIIIIIIIII
DB 147 QASSTYAPQP 156
RESULT 11
AAE12934
ID AAE12934 standard; Protein; 639 AA.
XX
AC AAE12934;
XX
DT 15-JAN-2002 (first entry)
XX
DE Dactylium dendroides galactose oxidase GO.05h1C variant.
XX
KM Galactose oxidase; GO; paper making; aldehyde synthesis; detection;
KM pulp biobleaching; colorectal adenocarcinoma; dental plaque; mutant;
KM mutein; variant.
XX
OS Dactylium dendroides.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 195 /note= "Wild type Gly substituted with Glu"
FT
FT Misc-difference 553 /note= "Wild type Ser substituted with Cys"
FT
XX
PN WO200162938-A2.
XX
PD 30-AUG-2001.
XX
PF 21-FEB-2001; 2001WO-US05732.
XX
PR 25-FEB-2000; 2000US-185001P.
PR 14-FEB-2001; 2001US-0782906.
XX
PA (HERC) HERCULES INC.
XX
PI Maffia AM, Delagrave S, Murphy DJ, Rittenhouse Pruss J, Bylina E;
PI Coleman WJ;
XX
DR WPI; 2001-648219/74.
XX
PT Novel variant galactose oxidase enzyme for oxidizing a compound such as
PT guar which is useful in paper making, for enzymatic synthesis of
PT aldehydes and pulp biobleaching, has at least one substituted amino
PT acid -

```
XX
PS Example 1; Page -: 65pp; English.
XX
CC The invention relates to variant galactose oxidase (VGO) proteins and
CC their CDNA molecules. Galactose oxidases are useful for oxidising
CC D-galactose side chains of many complex compounds such as guar, major
CC glycolipid of human red cells. GO is useful in paper making processes,
CC generation of H202 in situ, enzymatic synthesis of aldehydes, pulp
CC biobleaching, the use of GO-Schiff's reagent for early detection and
CC prognosis in human colorectal adenocarcinoma and use of GO-glucan
CC binding domain fusion proteins as targeting inhibitors of dental
CC plaque bacteria. GO is useful as a research tool for identification,
CC characterisation and purification of interacting regulatory proteins.
CC The present sequence is Dactylium dendroides galactose oxidase (GO)
CC variant.
CC Note: The present sequence is not shown in the specification but is
CC derived from Dactylium dendroides galactose oxidase SEQ ID NO: 2
CC shown in page 10 of the specification.
XX
SQ Sequence 639 AA;
XX
Query Match 91.2%; Score 52; DB 22; Length 639;
Best Local Similarity 100.0%; Pred. No. 0.82;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 QASSTYRPOP 10
Db 147 QASSTYRPOP 156
|||||
RESULT 12
AAE12935 standard; Protein; 639 AA.
XX
AC AAE12935;
XX
DT 15-JAN-2002 (first entry)
XX
DE Dactylium dendroides galactose oxidase GO.1hlc variant.
XX
KW Galactose oxidase; GO; paper making; aldehyde synthesis; detection;
KW pulp biobleaching; colorectal adenocarcinoma; dental plaque; mutant;
KW muten; variant.
XX
OS Dactylium dendroides.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT MISC-difference 6 /note= "Wild type Gly substituted with Arg"
FT WO200162938-A2.
XX
PN 30-AUG-2001.
XX
PD 21-FEB-2001; 2001WO-US05732.
XX
PR 25-FEB-2000; 2000US-185001P.
PR 14-FEB-2001; 2001US-0782906.
XX
PA (HERC ) HERCULES INC.
XX
PI Maffia AM, Delagrave S, Murphy DJ, Rittenhouse Pruss J, Bylina E;
PI Coleman WJ;
XX
DR WPI; 2001-648219/74.
XX
PT Novel variant galactose oxidase enzyme for oxidizing a compound such as
PT guar which is useful in paper making, for enzymatic synthesis of
PT aldehydes and pulp biobleaching, has at least one substituted amino
PT acid
XX
PS Example 1; Page -: 65pp; English.
```

```
XX
CC The invention relates to variant galactose oxidase (VGO) proteins and
CC their CDNA molecules. Galactose oxidases are useful for oxidising
CC D-galactose side chains of many complex compounds such as guar, major
CC glycolipid of human red cells. GO is useful in paper making processes,
CC generation of H202 in situ, enzymatic synthesis of aldehydes, pulp
CC biobleaching, the use of GO-Schiff's reagent for early detection and
CC prognosis in human colorectal adenocarcinoma and use of GO-glucan
CC binding domain fusion proteins as targeting inhibitors of dental
CC plaque bacteria. GO is useful as a research tool for identification,
CC characterisation and purification of interacting regulatory proteins.
CC The present sequence is Dactylium dendroides galactose oxidase (GO)
CC variant.
CC Note: The present sequence is not shown in the specification but is
CC derived from Dactylium dendroides galactose oxidase SEQ ID NO: 2
CC shown in page 10 of the specification.
XX
SQ Sequence 639 AA;
XX
Query Match 91.2%; Score 52; DB 22; Length 639;
Best Local Similarity 100.0%; Pred. No. 0.82;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 QASSTYRPOP 10
Db 147 QASSTYRPOP 156
|||||
RESULT 13
AAE12936 standard; Protein; 639 AA.
XX
AC AAE12936;
XX
DT 15-JAN-2002 (first entry)
XX
DE Dactylium dendroides galactose oxidase GO8.1h1a variant.
XX
KW Galactose oxidase; GO; paper making; aldehyde synthesis; detection;
KW pulp biobleaching; colorectal adenocarcinoma; dental plaque; mutant;
KW muten; variant.
XX
OS Dactylium dendroides.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT MISC-difference 238 /note= "Wild type Gln substituted with Leu"
FT MISC-difference 342 /note= "Wild type Lys substituted with Glu"
FT MISC-difference 383 /note= "Wild type Cys substituted with Ser"
FT MISC-difference 436 /note= "Wild type Tyr substituted with His"
FT MISC-difference 494 /note= "Wild type Val substituted with Ala"
FT WO200162938-A2.
XX
PN 30-AUG-2001.
XX
PD 21-FEB-2001; 2001WO-US05732.
XX
PR 25-FEB-2000; 2000US-185001P.
PR 14-FEB-2001; 2001US-0782906.
XX
PA (HERC ) HERCULES INC.
XX
PI Maffia AM, Delagrave S, Murphy DJ, Rittenhouse Pruss J, Bylina E;
PI Coleman WJ;
XX
DR WPI; 2001-648219/74.
XX
```

PT Novel variant galactose oxidase enzyme for oxidizing a compound such as
PT guar which is useful in paper making, for enzymatic synthesis of
PT aldehydes and pulp biobleaching, has at least one substituted amino
PT acid -
XX
XX Example 1; Page -: 65pp; English.
XX
CC The invention relates to variant galactose oxidase (VGO) proteins and
CC their cDNA molecules. Galactose oxidases are useful for oxidizing
CC D-galactose side chains of many complex compounds such as guar, major
CC glycolipid of human red cells. GO is useful in paper making processes,
CC generation of H2O2 in situ, enzymatic synthesis of aldehydes, pulp
CC biobleaching, the use of GO-Schiff's reagent for early detection and
CC prognosis in human colorectal adenocarcinoma and use of GO-glucan
CC binding domain fusion proteins as targeting inhibitors of dental
CC plaque bacteria. GO is useful as a research tool for identification,
CC characterisation and purification of interacting regulatory proteins.
CC The present sequence is Dactylium dendroides galactose oxidase (GO)
CC variant.
CC Note: The present sequence is not shown in the specification but is
CC derived from Dactylium dendroides galactose oxidase SEQ ID NO: 2
CC shown in page 10 of the specification.
XX
SQ Sequence 639 AA:

Query Match 91.2%; Score 52; DB 22; Length 639;
Best Local Similarity 100.0%; Pred. No. 0.82;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QASSTAPOP 10
|||||
DB 147 QASSTAPOP 156

RESULT 14
AAEL2937
ID AAEL2937 standard; Protein: 639 AA.
XX
XX AAEL2937;
XX
DT 15-JAN-2002 (first entry)
XX
DE Dactylium dendroides galactose oxidase G08-1h2A variant.
XX
KM Galactose oxidase; GO; paper making; aldehyde synthesis; detection;
KM pulp biobleaching; colorectal adenocarcinoma; dental plaque; mutant;
KM muten; variant.
XX
OS Dactylium dendroides.
OS Synthetic.
XX
XX Key Location/Qualifiers
FH Misc-difference 427
FT /note= "Wild type Asn substituted with Thr"
FT Misc-difference 436
FT /note= "Wild type Tyr substituted with His"
FT Misc-difference 494
FT /note= "Wild type Val substituted with Ala"
XX
XX WO200162938-A2.
XX
XX PD 30-AUG-2001.
XX
XX 21-FEB-2001; 2001WO-US05732.
XX
XX 25-FEB-2000; 2000US-185001P.
XX 14-FEB-2001; 2001US-0782906.
XX
XX (HERC) HERCULES INC.
XX
XX Maffia AM, Delagrave S, Murphy DJ, Rittenhouse Pruss J, Bylina E;
XX Coleman WJ;
XX

DR WPI; 2001-648219/74.
XX
XX Novel variant galactose oxidase enzyme for oxidizing a compound such as
PT guar which is useful in paper making, for enzymatic synthesis of
PT aldehydes and pulp biobleaching, has at least one substituted amino
PT acid -
XX
XX Example 1; Page -: 65pp; English.
XX
CC The invention relates to variant galactose oxidase (VGO) proteins and
CC their cDNA molecules. Galactose oxidases are useful for oxidizing
CC D-galactose side chains of many complex compounds such as guar, major
CC glycolipid of human red cells. GO is useful in paper making processes,
CC generation of H2O2 in situ, enzymatic synthesis of aldehydes, pulp
CC biobleaching, the use of GO-Schiff's reagent for early detection and
CC prognosis in human colorectal adenocarcinoma and use of GO-glucan
CC binding domain fusion proteins as targeting inhibitors of dental
CC plaque bacteria. GO is useful as a research tool for identification,
CC characterisation and purification of interacting regulatory proteins.
CC The present sequence is Dactylium dendroides galactose oxidase (GO)
CC variant.
CC Note: The present sequence is not shown in the specification but is
CC derived from Dactylium dendroides galactose oxidase SEQ ID NO: 2
CC shown in page 10 of the specification.
XX
SQ Sequence 639 AA:

Query Match 91.2%; Score 52; DB 22; Length 639;
Best Local Similarity 100.0%; Pred. No. 0.82;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QASSTAPOP 10
|||||
DB 147 QASSTAPOP 156

RESULT 15
AAEL2939
ID AAEL2939 standard; Protein: 639 AA.
XX
XX AAEL2939;
XX
DT 15-JAN-2002 (first entry)
XX
DE Dactylium dendroides galactose oxidase G08-1h4A variant.
XX
KM Galactose oxidase; GO; paper making; aldehyde synthesis; detection;
KM pulp biobleaching; colorectal adenocarcinoma; dental plaque; mutant;
KM muten; variant.
XX
OS Dactylium dendroides.
OS Synthetic.
XX
XX Key Location/Qualifiers
FH Misc-difference 63
FT /note= "Wild type Gln substituted with Lys"
FT Misc-difference 63
FT /note= "Wild type Gln substituted with Lys"
XX
XX WO200162938-A2.
XX
XX PD 30-AUG-2001.
XX
XX 21-FEB-2001; 2001WO-US05732.
XX
XX 25-FEB-2000; 2000US-185001P.
XX 14-FEB-2001; 2001US-0782906.
XX
XX (HERC) HERCULES INC.
XX
XX Maffia AM, Delagrave S, Murphy DJ, Rittenhouse Pruss J, Bylina E;
XX Coleman WJ;
XX WPI; 2001-648219/74.
XX

PT Novel variant galactose oxidase enzyme for oxidizing a compound such as
 PT guar which is useful in paper making, for enzymatic synthesis of
 PT aldehydes and pulp biobleaching, has at least one substituted amino
 PT acid

XX
 PS Example 1; Page -; 65pp; English.

XX
 CC The invention relates to variant galactose oxidase (VGO) proteins and
 CC their cDNA molecules. Galactose oxidases are useful for oxidizing
 CC D-galactose side chains of many complex compounds such as guar, major
 CC glycolipid of human red cells. GO is useful in paper making processes,
 CC generation of H2O2 in situ, enzymatic synthesis of aldehydes, pulp
 CC biobleaching, the use of GO-Schiff's reagent for early detection and
 CC prognosis in human colorectal adenocarcinoma and use of GO-glucan
 CC binding domain fusion proteins as targeting inhibitors of dental
 CC plaque bacteria. GO is useful as a research tool for identification,
 CC characterisation and purification of interacting regulatory proteins.
 CC The present sequence is Dactylium dendroides galactose oxidase (GO)
 CC variant.
 CC Note: The present sequence is not shown in the specification but is
 CC derived from Dactylium dendroides galactose oxidase SEQ ID NO: 2
 CC shown in page 10 of the specification.
 CC
 XX

SQ Sequence 639 AA;

Query Match 91.2%; Score 52; DB 22; Length 639;
 Best Local Similarity 100.0%; Pred. NO. 0.82;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QASSTAPOP 10
 |||||
 DB 147 QASSTAPOP 156

Search completed: January 13, 2003, 09:55:41
 Job time : 26.1026 secs

Mon Jan 13 10:35:14 2003

us-09-554-941-2.rspt

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 13, 2003, 09:53:20 : Search time 41.7436 Seconds
(without alignments)
54.296 Million cell updates/sec

Title: US-09-554-941-2
Perfect score: 57
Sequence: 1 QASSTAPQPP 11

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 671580 seqs, 206047115 residues 671580

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: SP archaea:*
2: SP bacteria:*
3: SP fungi:*
4: SP human:*
5: SP invertebrate:*
6: SP mammal:*
7: SP mhc:*
8: SP organelle:*
9: SP phage:*
10: SP plant:*
11: SP rodent:*
12: SP virus:*
13: SP vertebrate:*
14: SP unclassified:*
15: SP virus:*
16: SP bacteriophage:*
17: SP archaea:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	40	70.2	533	5 Q8WP81	O8WP81 caenorhabdi
2	39	68.4	224	16 Q9RU88	O9RU88 delinococcus
3	39	68.4	587	13 Q98SL7	O98SL7 xenopus lae
4	39	68.4	592	13 Q98SL6	O98SL6 xenopus lae
5	39	68.4	906	3 Q9P6N4	O9P6N4 schizosacch
6	38	66.7	124	12 Q9OH57	O9OH57 gallid herp
7	38	66.7	295	10 Q949H3	O949H3 nevaa brasi
8	38	66.7	859	11 Q9YDM6	O9YDM6 mus musculu
9	38	66.7	1004	4 P78354	P78354 homo sapien
10	37	64.9	90	16 Q9RWZ8	O9RWZ8 delinococcus
11	37	64.9	112	16 Q8YB7	O8YB7 brucea me
12	37	64.9	142	3 Q9U044	Q9U044 schizosacch
13	37	64.9	422	10 Q9XE77	Q9XE77 sorghum bic
14	37	64.9	583	5 Q61711	Q61711 caenorhabdi
15	37	64.9	1465	3 Q9P5N0	Q9P5N0 schizosacch
16	36	63.2	74	12 Q9IHS8	Q9IHS8 tt virus. o

ALIGNMENTS

17	36	63.2	139	2 Q54919	Q54919 streptococc
18	36	63.2	187	5 Q9VV20	Q9VV20 drosophila
19	36	63.2	231	2 Q9XC3	Q9XC3 streptococc
20	36	63.2	326	5 Q62010	Q62010 ceratilis c
21	36	63.2	329	10 Q9SM15	Q9SM15 zea mays (m
22	36	63.2	482	16 Q8YX0	Q8YX0 brucea me
23	36	63.2	546	16 Q9A911	Q9A911 caulobacter
24	36	63.2	681	5 Q9X0S8	Q9X0S8 caenorhabdi
25	36	63.2	781	16 Q9ZK65	Q9ZK65 thizobium m
26	36	63.2	1082	11 Q55201	Q55201 mus musculu
27	36	63.2	1084	13 Q9DDT5	Q9DDT5 brachydantio
28	36	63.2	1087	4 Q43279	Q43279 homo sapien
29	36	63.2	1087	4 Q96639	Q96639 homo sapien
30	36	63.2	1087	4 Q00267	Q00267 homo sapien
31	36	63.2	2658	13 Q9DWE0	Q9DWE0 gallus gall
32	36	63.2	4060	12 Q9IHZ8	Q9IHZ8 gill-assoc
33	35	61.4	115	11 Q8VHX8	Q8VHX8 mus musculu
34	35	61.4	153	12 Q70700	Q70700 human calic
35	35	61.4	201	10 Q9FE62	Q9FE62 arabidopsis
36	35	61.4	249	4 Q9H8H7	Q9H8H7 homo sapien
37	35	61.4	315	16 Q9L1T5	Q9L1T5 streptomyce
38	35	61.4	368	5 Q9E25	Q9E25 drosophila
39	35	61.4	369	5 Q9E25	Q9E25 drosophila
40	35	61.4	369	5 Q8T390	Q8T390 drosophila
41	35	61.4	420	5 Q8S747	Q8S747 drosophila
42	35	61.4	435	5 Q18380	Q18380 drosophila
43	35	61.4	435	5 Q9VB08	Q9VB08 drosophila
44	35	61.4	446	5 Q9VZC2	Q9VZC2 drosophila
45	35	61.4	483	16 Q9PEB1	Q9PEB1 xylella fas

RESULT 1
ID Q8WP81 PRELIMINARY: PRT: 533 AA.
AC Q8WP81
DT 01-MAR-2002 (TREMURel. 20, Created)
DT 01-MAR-2002 (TREMURel. 20, Last sequence update)
DT 01-JUN-2002 (TREMURel. 21, Last annotation update)
DE F13B12.6 protein.
GN F13B12.6.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Bardsall S.C.;
RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA MDLINF-99069613; PubMed=9851916;
RX none;
RT "genome sequence of the nematode C.elegans: A platform for
investigating biology."
RL Science 282:2012-2018(1998).
RN [3]
RP SEQUENCE FROM N.A.
RA Wild A.;
RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL: Z70686; CAD21656.1; JOINED.
DR EMBL: Z70683; CAD21626.1; JOINED.
DR EMBL: Z70686; CAD21626.1; JOINED.
DR EMBL: Z70686; IPR000980; SH2.
DR InterPro: IPR0017; SH2; 1.
DR Pfam: PF00017; SH2; 1.
DR ProDom: PD00093; SH2; 1.
DR SMART: SM00252; SH2; 1.
DR PROSITE: PS50001; SH2; 1.
SQ SEQUENCE 533 AA: 60897 MW: C244CF1884D08FF4 CRC64:

Query Match 70.2%; Score 40; DB 5; Length 533;
Best Local Similarity 77.8%; Pred. No. 26;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 S5YTAPOP 11
DB 185 S5YTAPOP 193

RESULT 2

Q98088 PRELIMINARY; PRT; 224 AA.
AC Q98088;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DE MADH dehydrogenase I, C subunit.
GN DRI504.
OS Deinococcus radiodurans.
OC Bacteria; Thermus/Deinococcus group; Deinococci; Deinococcales;
OC Deinococcaceae; Deinococcus.
RN NCBI_TaxID=1299;
RP SEQUENCE FROM N.A.
RC STRAIN-R1;
RX MEDLINE=20036896; PubMed=10567266;
RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
RA Dodson R.J., Haft D.H., Gwin M.L., Nelson W.C., Richardson D.L.,
RA Moritz K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
RA Vamathevan J.J., Lam P., McDonald L., Ueberlack T., Zaleski C.,
RA Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
RA Fraser C.M.;
RT "Genome sequence of the radioresistant bacterium Deinococcus
RT radiodurans R1";
RL Science 286:1571-1577 (1999).
DR EMBL: AE001994; AAF11071.1;
DR TIGR: DR1504;
DR InterPro: IPR001268; Complex1_30K.
DR Pfam: PF00329; Complex1_30Kd.1.
DR ProDom: PD001581; Complex1_30K.1.
DR PROSITE: PS00542; Complex1_30K.1.
KW Complete proteome.
SQ SEQUENCE 224 AA; 25032 MW; 05CC3BDF11281B35 CRC64;

Query Match 68.4%; Score 39; DB 16; Length 224;
Best Local Similarity 77.8%; Pred. No. 15;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 S5YTAPOP 11
DB 79 S5YTAPOP 87

RESULT 3

Q98SL7 PRELIMINARY; PRT; 587 AA.
AC Q98SL7;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DE EYES absent-1 alpha.
GN Xenopus laevis (African clawed frog).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
OC Xenopodinae; Xenopus.
RN NCBI_TaxID=8355;
RP SEQUENCE FROM N.A.
RC MEDLINE=21233041; PubMed=1135132;
RA David R., Ahrens K., Wedlich D., Schlosser G.;
RT "Xenopus Eyal demarcates all neurogenic placodes as well as migrating
RT hypaxial muscle precursors.";

RL Mech. Dev. 103:189-192(2001).
DR EMBL: AF352028; AAK31354.1;
DR InterPro: IPR001454; Hlganase/hydrlase.
DR Pfam: PF00702; Hydrolase.1.
SQ SEQUENCE 587 AA; 64142 MW; 5EE9813F6075ADE4 CRC64;

Query Match 68.4%; Score 39; DB 13; Length 587;
Best Local Similarity 87.5%; Pred. No. 44;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 S5YTAPOP 10
DB 162 S5YTAPOP 169

RESULT 4

Q98SL6 PRELIMINARY; PRT; 592 AA.
AC Q98SL6;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DE EYES absent-1 beta.
GN Xenopus laevis (African clawed frog).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
OC Xenopodinae; Xenopus.
RN NCBI_TaxID=8355;
RP SEQUENCE FROM N.A.
RC MEDLINE=21233041; PubMed=1135132;
RA David R., Ahrens K., Wedlich D., Schlosser G.;
RT "Xenopus Eyal demarcates all neurogenic placodes as well as migrating
RT hypaxial muscle precursors.";
RL Mech. Dev. 103:189-192(2001).
DR EMBL: AF352028; AAK31355.1;
DR InterPro: IPR001454; Hlganase/hydrlase.
DR Pfam: PF00702; Hydrolase.1.
SQ SEQUENCE 592 AA; 64640 MW; 84BE9396231B2C6 CRC64;

Query Match 68.4%; Score 39; DB 13; Length 592;
Best Local Similarity 87.5%; Pred. No. 44;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 S5YTAPOP 10
DB 167 S5YTAPOP 174

RESULT 5

Q9P6N4 PRELIMINARY; PRT; 906 AA.
AC Q9P6N4;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DE Putative vacuolar biogenesis protein, yeast endi homolog.
GN SPAC823.12.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC NCB1_TaxID=4896;
RN NCBI_TaxID=4896;
RP SEQUENCE FROM N.A.
RC STRAIN=972H-;
RA Voetkært G., Aert R., Wood V., Barrell B.G., Rajandream M.A.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AL355013; CAB90157.1;
DR InterPro: IPR000547; Clathrin_repeat.
DR Pfam: PF00637; Clathrin.1.
DR SMART: SM00299; CLH.1.

DR SMART; SM00184; RING: 1.
SQ SEQUENCE 906 AA; 102523 MW; 74F8B5B4440F3C80 CRC64;
Query Match 68.4%; Score 39; DB 3; Length 906;
Best Local Similarity 87.5%; Pred. No. 71;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 SYTAPOPO 11
|||||
Db 630 SYTAPOPO 637

RESULT 6
ID 090H57 PRELIMINARY: PRT: 124 AA.

AC 090H57; 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-MAY-2000 (TREMBlrel. 13, Last annotation update)
DE UL35 homolog.
OS Gallid herpesvirus 1.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Infectious laryngotracheitis-like viruses.
OX NCBI_TaxID=10386;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN-SA-2;
RA Johnson M.A.;
RT "Sequence of the infectious laryngotracheitis virus (SA-2 strain)
RT unique long region UL28 to UL43."
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF168792; AAD56207.1; -
SQ SEQUENCE 124 AA; 14267 MW; FE67191F6E3C7DA CRC64;
Query Match 66.7%; Score 38; DB 12; Length 124;
Best Local Similarity 54.5%; Pred. No. 12;
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 QASSTAPPO 11
|:|:|:|
Db 4 QSNHSTPOPO 14

RESULT 7
ID 0949H3 PRELIMINARY: PRT: 295 AA.
AC 0949H3; 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Putative class I chitinase (Fragment).
DE RO30.
GN Hevea brasiliensis (Para rubber tree).
OS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucosids I; Malpighiales; Euphorbiaceae; Hevea.
OX NCBI_TaxID=3981;
RN (1)
RP SEQUENCE FROM N.A.
RA O'Riordan G., Godric-Cavar J., Radauer C., Hoffmann-Sommergruber K.,
RA Scheiner O., Epner C., Breiteneder H.;
RT "Cloning and expression of a class I chitinase from latex, Hevea
RT brasiliensis."
RT Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
RL EMBL: AJ238579; CAC42881.1; -
DR InterPro: IPR001002; Chitin_binding_1.
DR InterPro: IPR000726; Glyco_hydro_19.
DR Pfam: PF00187; Chitin_bind_1;
DR Pfam: PF00182; Glyco_hydro_19; 1.
DR ProDom: PD000574; Glyco_hydro_19; 1.
DR PROSITE: PS00773; CHITINASE_19_1; UNKNOWN_1.
DR PROSITE: PS00774; CHITINASE_19_2; UNKNOWN_1.
KW Chitin-binding.

FT NON TER 1 1
SQ SEQUENCE 295 AA; 31647 MW; 6B748B2B3BFCF96 CRC64;
Query Match 66.7%; Score 38; DB 10; Length 295;
Best Local Similarity 70.0%; Pred. No. 31;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 QASSTAPPO 10
|||||
Db 142 QASSTAPPO 151

RESULT 8
ID 08VDM6 PRELIMINARY: PRT: 859 AA.

AC 08VDM6; 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Similar to E1B-55 kDa-associated protein 5.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN (1)
RP SEQUENCE FROM N.A.
RA Strauberg R.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC021506; AAN1506.1; -
DR InterPro: IPR003034; SAP.
DR InterPro: IPR003878; SPRY domain.
DR InterPro: IPR003877; SPRY_receptor.
DR Pfam: PF02037; SAP; 1.
DR Pfam: PF00622; SPRY; 1.
DR SMART: SM00513; SAP; 1.
DR SMART: SM00449; SPRY; 1.
SQ SEQUENCE 859 AA; 96002 MW; 197328B681DF260E CRC64;
Query Match 66.7%; Score 38; DB 11; Length 859;
Best Local Similarity 87.5%; Pred. No. 1e+02; 1; Indels 0; Gaps 0;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 SSYTAPOPO 10
|||||
Db 751 SSYTAPOPO 758

RESULT 9
ID P78364 PRELIMINARY: PRT: 1004 AA.
AC P78364; 01-MAY-1997 (TREMBlrel. 03, Created)
DT 01-MAY-1997 (TREMBlrel. 03, Last sequence update)
DT 01-MAY-1997 (TREMBlrel. 19, Last annotation update)
DE Polyhomeotic 1 homolog (Fragment).
DE HP1.
GN Homo sapiens (Human).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN (1)
RP SEQUENCE FROM N.A.
RA MEDLINE-97220024; PubMed-9121482;
RX Gunster M.J., Satijn D.P., Hamer K.M., den Blaauwen J.L.,
RA de Bruijn D., Alkema M.J., van Loonhuizen M., van Driel R.,
RT "Identification and characterization of interactions between the
RT vertebrate polycomb-group protein Bmi1 and human homologs of
RT polyhomeotic."
RL Mol. Cell. Biol. 17:2326-2335(1997).
DR EMBL: U89277; AAC51168.1; -
DR InterPro: IPR001660; SAM.
DR Pfam: PF00536; SAM; 1.
DR SMART: SM00454; SAM; 1.

FT NON-TER 1004 1004
SQ SEQUENCE 1004 AA; 105479 MW; D2A61904C7D32EB1 CRC64;
Query Match
Best Local Similarity 66.7%; Score 38; DB 4; Length 1004;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 1 QASSTAPQ 10
DB 441 QATLTAPQ 450

RESULT 10

Q9RWZ8 PRELIMINARY; PRT; 90 AA.
AC Q9RWZ8
DT 01-MAY-2000 (TREMblrel. 13, Created)
DE 01-MAY-2000 (TREMblrel. 13, Last sequence update)
DE 01-MAY-2000 (TREMblrel. 20, Last annotation update)
GN Hypothetical protein DR0517.
OS Deinococcus radiodurans.
OC Bacteria; Thermus/Deinococcus group; Deinococci; Deinococcales;
OC Deinococcaceae; Deinococcus.
OX NCBI_TaxID=1299;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=R1;
RX MEDLINE=20036896; PubMed=10567266;
RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
Moffatt K.S., Qin H., Jiang L., Nelson W.C., Richardson D.L.,
Vamathevan J.J., Lam P., McDonald L., Uitterback T., Zaleski C.,
Ketchum K.A., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
Fraser C.M.;
RT "Genome sequence of the radioresistant bacterium Deinococcus
radiodurans R1.";
RL Science 286:1571-1577(1999).
DR EMBL: AE001910; AAF10097.1; -;
DR TIGR: DR0517; -;
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 90 AA; 8948 MW; F6DB1AB9F1PDE5D CRC64;
Query Match
Best Local Similarity 64.9%; Score 37; DB 16; Length 90;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 2 ASSYTAPOQ 11
DB 36 ASSYTAPOQ 45

RESULT 11

Q8YBW7 PRELIMINARY; PRT; 112 AA.
AC Q8YBW7
DT 01-MAR-2002 (TREMblrel. 20, Created)
DE 01-MAR-2002 (TREMblrel. 20, Last sequence update)
DE 01-JUN-2002 (TREMblrel. 21, Last annotation update)
GN Potassium efflux system protein PHAC.
OS Brucella melitensis.
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Brucellaceae; Brucella.
OX NCBI_TaxID=29459;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=16M / ATCC 23456 / BIOTYPE 1;
RX MEDLINE=20020109; PubMed=11756688;
RA Delvecchio V.G., Kapatal V., Redkar R.J., Patra G., Mujer C., Los T.,
Ivanova N., Anderson T., Bhattacharya A., Lykidis A., Reznik G.,
Jablonski L., Larsen N., D'Souza M., Bernal A., Mazur M., Goltsman E.,

RA Selkov E., Elizer P.H., Hagius S., O'Callaghan D., Letesson J.-J.,
Haselkorn R., Kyriades N., Overbeek R.;
RT "The genome sequence of the facultative intracellular pathogen
Brucella melitensis.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:443-448(2002).
DR EMBL: AE009711; AAL54011.1; -;
DR InterPro: IPR002757; DUF67.
DR Pfam: PF01898; DUF67.1.
DR ProDom: PD006097; DUF67.1.
KW Complete proteome.
SQ SEQUENCE 112 AA; 12010 MW; D386BE18F5EE81B CRC64;
Query Match
Best Local Similarity 64.9%; Score 37; DB 16; Length 112;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
QY 1 QASSTAPQ 11
DB 64 QAAQYTDYVPQ 74

RESULT 12

Q9U044 PRELIMINARY; PRT; 142 AA.
AC Q9U044
DT 01-MAY-2000 (TREMblrel. 13, Created)
DE 01-MAY-2000 (TREMblrel. 13, Last sequence update)
DE 01-JUN-2002 (TREMblrel. 21, Last annotation update)
DE Hypothetical 16.6 kDa protein (Fragment)
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=968 H90;
RX MEDLINE=20223868; PubMed=10759889;
RA Ding D.O., Tomita Y., Yamamoto A., Chikashige Y., Haraguchi T.,
Hirooka Y.;
RT "Large-scale screening of intracellular protein localization in living
fission yeast cells by the use of a GFP-fusion genomic DNA library.";
RL Genes Cells 5:169-190(2000).
DR EMBL: AB027821; BA87125.1; -;
DR InterPro: IPR001140; ABCtransprtm.
DR Pfam: PF00664; ABC_membrane; 1.
KW Hypothetical protein.
FT NON-TER 142 142
SQ SEQUENCE 142 AA; 16561 MW; 8296B1B8BD0B5C4F CRC64;
Query Match
Best Local Similarity 64.9%; Score 37; DB 3; Length 142;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 3 SSTYAPQ 11
DB 73 SSTYSPNE 81

RESULT 13

Q9XE77 PRELIMINARY; PRT; 422 AA.
AC Q9XE77
DT 01-NOV-1999 (TREMblrel. 12, Created)
DT 01-NOV-1999 (TREMblrel. 12, Last sequence update)
DE Patatin-like protein.
OS Sorghum bicolor (Sorghum) (Sorghum vulgare).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
OC Panicoideae; Andropogoneae; Sorghum.
OX NCBI_TaxID=4558;

Mon Jan 13 10:35:14 2003

Search completed: January 13, 2003, 09:59:45
Job time : 44.7436 secs

us-09-554-941-2.rspt

GenCore version 5.1.3
2003 CompuGen Ltd.

using sw model
+time 11 seconds

A.L.4.10

09-554-941-2

perfect sequence:
1 QABD
0.5

112892

total number of files: 0
length: 0

Maximum existing	Match first	45 summaries
1	1	1
2	2	2
3	3	3
4	4	4
5	5	5
6	6	6
7	7	7
8	8	8
9	9	9
10	10	10
11	11	11
12	12	12
13	13	13
14	14	14
15	15	15
16	16	16
17	17	17
18	18	18
19	19	19
20	20	20
21	21	21
22	22	22
23	23	23
24	24	24
25	25	25
26	26	26
27	27	27
28	28	28
29	29	29
30	30	30
31	31	31
32	32	32
33	33	33
34	34	34
35	35	35
36	36	36
37	37	37
38	38	38
39	39	39
40	40	40
41	41	41
42	42	42
43	43	43
44	44	44
45	45	45
46	46	46
47	47	47
48	48	48
49	49	49
50	50	50
51	51	51
52	52	52
53	53	53
54	54	54
55	55	55
56	56	56
57	57	57
58	58	58
59	59	59
60	60	60
61	61	61
62	62	62
63	63	63
64	64	64
65	65	65
66	66	66
67	67	67
68	68	68
69	69	69
70	70	70
71	71	71
72	72	72
73	73	73
74	74	74
75	75	75
76	76	76
77	77	77
78	78	78
79	79	79
80	80	80
81	81	81
82	82	82
83	83	83
84	84	84
85	85	85
86	86	86
87	87	87
88	88	88
89	89	89
90	90	90
91	91	91
92	92	92
93	93	93
94	94	94
95	95	95
96	96	96
97	97	97
98	98	98
99	99	99
100	100	100

Database :
is the number of results produced by the search engine
is equal to the total score distribution

SUMMARY

Query Match	Length	ID	Description
91.2	680	1	GAOA_DMCDE
64.9	423	1	MDRC_HUMAN
64.9	423	1	MDRC_MOUSE
63.2	384	1	STPR_STRGR
63.2	384	1	CP24_HUMAN
63.2	663	1	CP24_MOUSE
63.2	663	1	CASB_CAMDR
61.4	286	1	IF2M_MOUSE
61.4	286	1	IF2M_HUMAN
61.4	287	1	PKNA_STRGC
59.6	34	1	VE1_HCV54
57.9	175	1	CPG3_CANGG
57.9	175	1	FLJ3_HILPN
57.9	262	1	OMPY_CHIMP
57.9	297	1	GDPA_MOUSE
57.9	310	1	HALO_XANAU
57.9	332	1	LPKX_PSEEU
57.9	459	1	LDL1_MOUSE
57.9	459	1	SOX3_MOUSE
57.9	464	1	CHIT_MOUSE
57.9	551	1	ACES_BUNFA
57.9	581	1	VE1_MOUSE
57.9	595	1	DDX3_MOUSE
57.9	614	1	HMIA_DROME
57.9	706	1	FTSL_ASPOP
57.9	866	1	AREX_YEAST
57.9	926	1	UBR4_YEAST
57.9	956	1	VE1_MOUSE
57.9	982	1	LILO_CAEEL
57.9	1097	1	PRD_YEAST
57.9	1403	1	YCO3_MYCPN
56.1	221	1	GRPE_BACST
91.2	680	1	GAOA_DMCDE
64.9	423	1	MDRC_HUMAN
64.9	423	1	MDRC_MOUSE
63.2	384	1	STPR_STRGR
63.2	384	1	CP24_HUMAN
63.2	663	1	CP24_MOUSE
63.2	663	1	CASB_CAMDR
61.4	286	1	IF2M_MOUSE
61.4	286	1	IF2M_HUMAN
61.4	287	1	PKNA_STRGC
59.6	34	1	VE1_HCV54
57.9	175	1	CPG3_CANGG
57.9	175	1	FLJ3_HILPN
57.9	262	1	OMPY_CHIMP
57.9	297	1	GDPA_MOUSE
57.9	310	1	HALO_XANAU
57.9	332	1	LPKX_PSEEU
57.9	459	1	LDL1_MOUSE
57.9	459	1	SOX3_MOUSE
57.9	464	1	CHIT_MOUSE
57.9	551	1	ACES_BUNFA
57.9	581	1	VE1_MOUSE
57.9	595	1	DDX3_MOUSE
57.9	614	1	HMIA_DROME
57.9	706	1	FTSL_ASPOP
57.9	866	1	AREX_YEAST
57.9	926	1	UBR4_YEAST
57.9	956	1	VE1_MOUSE
57.9	982	1	LILO_CAEEL
57.9	1097	1	PRD_YEAST
57.9	1403	1	YCO3_MYCPN
56.1	221	1	GRPE_BACST
91.2	680	1	GAOA_DMCDE
64.9	423	1	MDRC_HUMAN
64.9	423	1	MDRC_MOUSE
63.2	384	1	STPR_STRGR
63.2	384	1	CP24_HUMAN
63.2	663	1	CP24_MOUSE
63.2	663	1	CASB_CAMDR
61.4	286	1	IF2M_MOUSE
61.4	286	1	IF2M_HUMAN
61.4	287	1	PKNA_STRGC
59.6	34	1	VE1_HCV54
57.9	175	1	CPG3_CANGG
57.9	175	1	FLJ3_HILPN
57.9	262	1	OMPY_CHIMP
57.9	297	1	GDPA_MOUSE
57.9	310	1	HALO_XANAU
57.9	332	1	LPKX_PSEEU
57.9	459	1	LDL1_MOUSE
57.9	459	1	SOX3_MOUSE
57.9	464	1	CHIT_MOUSE
57.9	551	1	ACES_BUNFA
57.9	581	1	VE1_MOUSE
57.9	595	1	DDX3_MOUSE
57.9	614	1	HMIA_DROME
57.9	706	1	FTSL_ASPOP
57.9	866	1	AREX_YEAST
57.9	926	1	UBR4_YEAST
57.9	956	1	VE1_MOUSE
57.9	982	1	LILO_CAEEL
57.9	1097	1	PRD_YEAST
57.9	1403	1	YCO3_MYCPN
56.1	221	1	GRPE_BACST
91.2	680	1	GAOA_DMCDE
64.9	423	1	MDRC_HUMAN
64.9	423	1	MDRC_MOUSE
63.2	384	1	STPR_STRGR
63.2	384	1	CP24_HUMAN
63.2	663	1	CP24_MOUSE
63.2	663	1	CASB_CAMDR
61.4	286	1	IF2M_MOUSE
61.4	286	1	IF2M_HUMAN
61.4	287	1	PKNA_STRGC
59.6	34	1	VE1_HCV54
57.9	175	1	CPG3_CANGG
57.9	175	1	FLJ3_HILPN
57.9	262	1	OMPY_CHIMP
57.9	297	1	GDPA_MOUSE
57.9	310	1	HALO_XANAU
57.9	332	1	LPKX_PSEEU
57.9	459	1	LDL1_MOUSE
57.9	459	1	SOX3_MOUSE
57.9	464	1	CHIT_MOUSE
57.9	551	1	ACES_BUNFA
57.9	581	1	VE1_MOUSE
57.9	595	1	DDX3_MOUSE
57.9	614	1	HMIA_DROME
57.9	706	1	FTSL_ASPOP
57.9	866	1	AREX_YEAST
57.9	926	1	UBR4_YEAST
57.9	956	1	VE1_MOUSE
57.9	982	1	LILO_CAEEL
57.9	1097	1	PRD_YEAST
57.9	1403	1	YCO3_MYCPN
56.1	221	1	GRPE_BACST
91.2	680	1	GAOA_DMCDE
64.9	423	1	MDRC_HUMAN
64.9	423	1	MDRC_MOUSE
63.2	384	1	STPR_STRGR
63.2	384	1	CP24_HUMAN
63.2	663	1	CP24_MOUSE
63.2	663	1	CASB_CAMDR
61.4	286	1	IF2M_MOUSE
61.4	286	1	IF2M_HUMAN
61.4	287	1	PKNA_STRGC
59.6	34	1	VE1_HCV54
57.9	175	1	CPG3_CANGG
57.9	175	1	FLJ3_HILPN
57.9	262	1	OMPY_CHIMP
57.9	297	1	GDPA_MOUSE
57.9	310	1	HALO_XANAU
57.9	332	1	LPKX_PSEEU
57.9	459	1	LDL1_MOUSE
57.9	459	1	SOX3_MOUSE
57.9	464	1	CHIT_MOUSE
57.9	551	1	ACES_BUNFA
57.9	581	1	VE1_MOUSE
57.9	595	1	DDX3_MOUSE
57.9	614	1	HMIA_DROME
57.9	706	1	FTSL_ASPOP
57.9	866	1	AREX_YEAST
57.9	926	1	UBR4_YEAST
57.9	956	1	VE1_MOUSE
57.9	982	1	LILO_CAEEL
57.9	1097	1	PRD_YEAST
57.9	1403	1	YCO3_MYCPN
56.1	221	1	GRPE_BACST
91.2	680	1	GAOA_DMCDE
64.9	423	1	MDRC_HUMAN
64.9	423	1	MDRC_MOUSE
63.2	384	1	STPR_STRGR
63.2	384	1	CP24_HUMAN
63.2	663	1	CP24_MOUSE
63.2	663	1	CASB_CAMDR
61.4	286	1	IF2M_MOUSE
61.4	286	1	IF2M_HUMAN
61.4	287	1	PKNA_STRGC
59.6	34	1	VE1_HCV54
57.9	175	1	CPG3_CANGG
57.9	175	1	FLJ3_HILPN
57.9	262	1	OMPY_CHIMP
57.9	297	1	GDPA_MOUSE
57.9	310	1	HALO_XANAU
57.9	332	1	LPKX_PSEEU
57.9	459	1	LDL1_MOUSE
57.9	459	1	SOX3_MOUSE
57.9	464	1	CHIT_MOUSE
57.9	551	1	ACES_BUNFA
57.9	581	1	VE1_MOUSE
57.9	595	1	DDX3_MOUSE
57.9	614	1	HMIA_DROME
57.9	706	1	FTSL_ASPOP
57.9	866	1	AREX_YEAST
57.9	926	1	UBR4_YEAST
57.9	956	1	VE1_MOUSE
57.9	982	1	LILO_CAEEL
57.9	1097	1	PRD_YEAST
57.9	1403	1	YCO3_MYCPN
56.1	221	1	GRPE_BACST
91.2	680	1	GAOA_DMCDE
64.9	423	1	MDRC_HUMAN
64.9	423	1	MDRC_MOUSE
63.2	384	1	STPR_STRGR
63.2	384	1	CP24_HUMAN
63.2	663	1	CP24_MOUSE
63.2	663	1	CASB_CAMDR
61.4	286	1	IF2M_MOUSE
61.4	286	1	IF2M_HUMAN
61.4	287	1	PKNA_STRGC
59.6	34	1	VE1_HCV54
57.9	175	1	CPG3_CANGG
57.9	175	1	FLJ3_HILPN
57.9	262	1	OMPY_CHIMP
57.9	297	1	GDPA_MOUSE
57.9	310	1	HALO_XANAU
57.9	332	1	LPKX_PSEEU
57.9	459	1	LDL1_MOUSE
57.9	459	1	SOX3_MOUSE
57.9	464	1	CHIT_MOUSE
57.9	551	1	ACES_BUNFA
57.9	581	1	VE1_MOUSE
57.9	595	1	DDX3_MOUSE
57.9	614	1	HMIA_DROME
57.9	706	1	FTSL_ASPOP
57.9	866	1	AREX_YEAST
57.9	926	1	UBR4_YEAST
57.9	956	1	VE1_MOUSE
57.9	982	1	LILO_CAEEL
57.9	1097	1	PRD_YEAST
57.9	1403	1	YCO3_MYCPN
56.1	221	1	GRPE_BACST
91.2	680	1	GAOA_DMCDE
64.9	423	1	MDRC_HUMAN
64.9	423	1	MDRC_MOUSE
63.2	384	1	STPR_STRGR
63.2	384	1	CP24_HUMAN
63.2	663	1	CP24_MOUSE
63.2	663	1	CASB_CAMDR
61.4	286	1	IF2M_MOUSE
61.4	286	1	IF2M_HUMAN
61.4	287	1	PKNA_STRGC
59.6	34	1	VE1_HCV54
57.9	175	1	CPG3_CANGG
57.9	175	1	FLJ3_HILPN
57.9	262	1	OMPY_CHIMP
57.9	297	1	GDPA_MOUSE
57.9	310	1	HALO_XANAU
57.9	332	1	LPKX_PSEEU
57.9	459	1	LDL1_MOUSE
57.9	459	1	SOX3_MOUSE
57.9	464	1	CHIT_MOUSE
57.9	551	1	ACES_BUNFA
57.9	581	1	VE1_MOUSE
57.9	595	1	DDX3_MOUSE
57.9	614	1	HMIA_DROME
57.9	706	1	FTSL_ASPOP
57.9	866	1	AREX_YEAST
57.9	926	1	UBR4_YEAST
57.9	956	1	VE1_MOUSE
57.9	982	1	LILO_CAEEL
57.9	1097	1	PRD_YEAST
57.9	1403	1	YCO3_MYCPN
56.1	221	1	GRPE_BACST
91.2	680	1	GAOA_DMCDE
64.9	423	1	MDRC_HUMAN
64.9	423	1	MDRC_MOUSE
63.2	384	1	STPR_STRGR
63.2	384	1	CP24_HUMAN
63.2	663	1	CP24_MOUSE
63.2	663	1	CASB_CAMDR
61.4	286	1	IF2M_MOUSE
61.4	286	1	IF2M_HUMAN
61.4	287	1	PKNA_STRGC
59.6	34	1	VE1_HCV54
57.9	175	1	CPG3_CANGG
57.9	175	1	FLJ3_HILPN
57.9	262	1	OMPY_CHIMP
57.9	297	1	GDPA_MOUSE
57.9	310	1	HALO_XANAU
57.9	332	1	LPKX_PSEEU
57.9	459	1	LDL1_MOUSE
57.9	459	1	SOX3_MOUSE
57.9	464	1	CHIT_MOUSE
57.9	551	1	ACES_BUNFA
57.9	581	1	VE1_MOUSE
57.9	595	1	DDX3_MOUSE
57.9	614	1	HMIA_DROME
57.9	706	1	FTSL_ASPOP
57.9	866	1	AREX_YEAST
57.9	926	1	UBR4_YEAST
57.9	956	1	VE1_MOUSE
57.9	982	1	LILO_CAEEL
57.9	1097	1	PRD_YEAST
57.9	1403	1	YCO3_MYCPN
56.1	221	1	GRPE_BACST
91.2	680	1	GAOA_DMCDE
64.9	423	1	MDRC_HUMAN
64.9	423	1	MDRC_MOUSE
63.2	384	1	STPR_STRGR
63.2	384	1	CP24_HUMAN
63.2	663	1	CP24_MOUSE
63.2	663	1	CASB_CAMDR
61.4	286	1	IF2M_MOUSE
61.4	286	1	IF2M_HUMAN
61.4	287	1	PKNA_STRGC
59.6	34	1	VE1_HCV54
57.9	175	1	CPG3_CANGG
57.9	175	1	FLJ3_HILPN
57.9	262	1	OMPY_CHIMP
57.9	297	1	GDPA_MOUSE
57.9	310	1	HALO_XANAU
57.9	332	1	LPKX_PSEEU
57.9	459	1	LDL1_MOUSE
57.9	459	1	SOX3_MOUSE
57.9	464	1	CHIT_MOUSE
57.9	551	1	ACES_BUNFA
57.9	581	1	VE1_MOUSE
57.9	595	1	DDX3_MOUSE
57.9	614	1	HMIA_DROME
57.9	706	1	FTSL_ASPOP
57.9	866	1	AREX_YEAST
57.9	926	1	UBR4_YEAST
57.9	956	1	VE1_MOUSE
57.9	982	1	LILO_CAEEL
57.9	1097	1	PRD_YEAST
57.9	1403	1	YCO3_MYCPN
56.1	221	1	GRPE_BACST
91.2	680	1	GAOA_DMCDE
64.9	423	1	MDRC_HUMAN
64.9	423	1	MDRC_MOUSE
63.2	384	1	STPR_STRGR
63.2	384	1	CP24_HUMAN
63.2	663	1	CP24_MOUSE
63.2	663	1	CASB_CAMDR
61.4	286	1	IF2M_MOUSE
61.4	286	1	IF2M_HUMAN
61.4	287	1	PKNA_STRGC
59.6	34	1	VE1_HCV54
57.9	175	1	CPG3_CANGG
57.9	175	1	FLJ3_HILPN
57.9	262	1	OMPY_CHIMP
57.9	297	1	GDPA_MOUSE
57.9	310	1	HALO_XANAU
57.9	332	1	LPKX_PSEEU
57.9	459	1	LDL1_MOUSE
57.9	459		

45	32	56.1	77.5	1	GURP_MYXA
44	32	56.1	77.5	1	GURP_DICDI
43	32	56.1	77.5	1	SVDI_MOUSE
42	32	56.1	66.3	1	W448_APRTH
41	32	56.1	66.3	1	W448_DEIRA
40	32	56.1	61.7	1	YC46_ODPSI
39	32	56.1	61.7	1	NC45_CANEL
38	32	56.1	57.8	1	YH35_CANEL
37	32	56.1	49.7	1	ADFI_CANEL
36	32	56.1	49.7	1	VGFR_RAPR
35	32	56.1	49.7	1	VGFR_RAPR
34	32	56.1	49.7	1	UYRB_AOUAE
33	32	56.1	49.7	1	UYRB_AOUAE
32	32	56.1	49.7	1	GCR_CAYMO
31	32	56.1	49.7	1	PINC_MOUSE
30	32	56.1	49.7	1	PINC_MOUSE

Q06916 myxococcus
P19188 dictyostell
Q35188 mus musculu
Q91924 arbidopsis
O91968 delnoccus
P49541 odontella s
Q17711 caenorhabdi
P46599 candida alb
P20156 ratus norv
O67708 aquifex aeo
P49115 cavia porce
P35831 mus musculu

RESULT 1
GAO#_DACDE STANDARD; PRT:
ID GAO#_DACDE
AC 001745; 1993 (Rel. 26, Created) (rel. sequence update)
AD 01-JUL-1993 (Rel. 26, Last annotation update)
DT 01-JUL-1993 (Rel. 40, Last annotation update) (GAO).
DR 16-OCT-2001 (Rel. 40, Last annotation update) (EC 1.1.3.9) (GAO).
DN Galactose oxidase precursor (EC 1.1.3.9) (GAO).
DE Dactylium dendroides (Cladobotryum dendroides); Sordariomycetes;
GC Dactylium dendroides (Cladobotryum dendroides);
OS Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Hypocreales; Hypocreaceae; Hypomyces.
NCI TaxID=5132;
[1] SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
MEDLINE=92235025; PubMed=1569070; Yadvav K.D.S., Keen J.N.,
McPherson M.J., Ogel Z.B., Stevens C.E., Yadvav K.D.S., Keen J.N.,
McPherson M.J., Knowles R.F.;
"galactose oxidase of Dactylium dendroides".
analysts";
J. Biol. Chem. 267:8146-8152(1992).
[2] X-RAY CRYSTALLOGRAPHY [1.7 ANGSTROMS].
X-RAY CRYSTALLOGRAPHY [1.7 ANGSTROMS]; Ogel Z.B., McPherson M.J.,
MEDLINE=91163641; PubMed=2002850 C.E., Ogel Z.B., McPherson M.J.,
Stevens P.F.;
ITO N., Phillips S.P., Knowles R.F.;
Keen J.N., Yadvav K.D.S.,
"Novel thioether bond revealed by a 1.7 A crystal structure of
galactose oxidase [1991]. THE STEREOSPECIFIC OXIDATION OF A BROAD RANGE
NATURE FUNCTION: CATALYZES THE STEREOSPECIFIC OXIDATION OF A BROAD RANGE
-1 FUNCTION: CATALYZES THE STEREOSPECIFIC OXIDATION OF A BROAD RANGE
OF PRIMARY ALCOHOL SUBSTRATES.
-1 CATALYTIC ACTIVITY: D-galactose + O(2) -> D-galacto-hexodialdose +
H2O(2). COPPER.
-1 COFACTOR: MONOMER.
-1 SUBUNIT: MONOMER.
-1 SUBCELLULAR LOCATION: Secreted.
-1 SIMILILARITY: CONTAINS 3 KEICH REPEATS.
CC -1 DATABASE: NMBE-WORTHINGTON-BIOCHEM.COM/MANUAL/G/GNO.HTML"
CC -1 WWW:"http://www.worthington-biochem.com/manual/g/gno.html".
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute Institute. There are no restrictions on way
CC the European Bioinformatics Institute. These are no restrictions on way
CC used by non-profit institutions as long as its content is in
CC modified and this statement agreement (see http://www.isb-sib.ch/announce,
CC entitles requires a licenseesib-sib.ch).
CC or send an email to license@sib-sib.ch).
CC
CC EMBL: M86819; AAA16228.1;
CC PDB: 1GOF; 31 JAN-94.
CC PDB: 1GOB; 31 JAN-94.
CC PDB: 1GOH; 31 JAN-94.
CC INTERPRO: IPRO01798; KEICH.3.
CC InterPro: IPRO01798; KEICH.3.
CC Pfam: PF01344; KEICHREPEAT.
CC PRINTS: PR00501; KEICHREPEAT.

Page 2

[illegible]

MDRC_MOUSE STANDARD: PRT: 423 AA.

RESULT 3

ID MDRC_MOUSE STANDARD: PRT: 423 AA.

AC 09JUN04; 09JUN05; 09JUN03;

DT 15-JUN-2002 (Rel. 41, Created)

DT 15-JUN-2002 (Rel. 41, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE WD-repeat protein 12 (YTM1 homolog).

GN WDR12.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OC NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Thymus;

RX MEDLINE=21686157; PubMed=11827460;

RA Nal B., Mohr E., Da Silva M.I., Tagett R., Navarro C., Carroll P.,

RA Depetris D., Verthuy C., Jordan B.R., Ferrier P.;

RT "Wdr12, a mouse gene encoding a novel WD-repeat protein with a

RT notchless-like amino-terminal domain."

RL Genomics 79:77-86(2002).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=129/SVJ X C57BL/6J;

RA Matsumoto S.;

RT "Mouse homologue of Saccharomyces cerevisiae YTM1".

RT Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.

RN [3]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6; TISSUE=Brain;

RA Osada N., Kusuda J., Tanuma R., Ito A., Hirata M., Sugano S.,

RA Hashimoto K.;

RT "Isolation of full-length cDNA clones from mouse brain cDNA library

RT made by oligo-capping method."

RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.

RN [4]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Testis, and Pancreas;

RX MEDLINE=21085660; PubMed=11217851;

RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishi Y.,

RA Aikawa T., Hata A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,

RA Aikawa K., Ikawa M., Nishi K., Kiyosawa H., Kondo S., Yamana K.,

RA Saito T., Ozaki Y., Gojodori T., Hono H., Kasukawa T., Saito R.,

RA Kaeda K., Matsuda H.A., Ashburner M., Batilov S., Casavant T.,

RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochia H.,

RA Knehl P., Lewis S., Matsuo Y., Nikolaou I., Pesole G., Quackenbush J.,

RA Schiml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,

RA Sakai K., Offord T., Furuno M., Aono H., Baladrelli R., Barch G.,

RA Blake J., Boffelli D., Bojunga N., Carlini P., Fujita M., Gariboldi M.,

RA Brownstein M.J., Bult C., Hofmann M., Hume D.A., Kamita M., Lee N.H.,

RA Gustinchik S., Hill D., Mashima M., Rodriguez I., Sakamoto N.,

RA Lyons P., Marchionni L., Mashima M., Rodriguez I., Sakamoto N.,

RA Norodone P., Ring B., Schenck C., Seya T., Shibata Y., Storch K.-F.,

RA Sasaki H., Sato K., Schenck C., Seya T., Shibata Y., Storch K.-F.,

RA Suzuki H., Toyooka K., Wang K.H., Weltz C., Whitaker C., Wilming L.,

RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,

RA Hayashizaki Y.;

RT "Functional annotation of a full-length mouse cDNA collection";

RL Nature 409:685-690(2001).

RN [5]

RP SEQUENCE FROM N.A.

RA Strausberg R.;

RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.

CC -1- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

```

CC EMBL: AY059431; AAL29680.1;
DR EMBL: AY059432; AAL29681.1;
DR EMBL: AF239765; AAF44683.1;
DR EMBL: AB041608; BAB5091.1;
DR EMBL: AK007500; BAB5072.1;
DR EMBL: AK012022; BAB27979.1;
DR EMBL: AK016611; BAB30336.1;
DR EMBL: BC004748; AAO04748.1;
DR MGD: MGI:1927241; MGI:12;
DR InterPro: IPR001680; WD40;
DR Pfam: PF00400; WD40; 7;
DR PRINTS: PR00320; GPROTEINRPT;
DR ProDom: PD000018; WD40; 2;
DR SMART: SM00320; WD40; 7;
DR PROSITE: PS00678; WD_REPEATS_1; 2;
DR PROSITE: PS50082; WD_REPEATS_2; 5;
DR PROSITE: PS50294; WD_REPEATS_REGION; 1;
KW Repeat; WD repeat.
FT REPEAT 99 137 WD 1;
FT REPEAT 138 180 WD 2;
FT REPEAT 187 226 WD 3;
FT REPEAT 255 293 WD 4;
FT REPEAT 295 334 WD 5;
FT REPEAT 340 380 WD 6;
FT REPEAT 422 442 WD 7;
FT REPEAT 499 518 WD 7;
FT CONFLICT 418 418 F -> S (IN REF. 2);
FT CONFLICT 418 418 T -> A (IN REF. 1);
SQ SEQUENCE 423 AA; 47346 MW; DF57952520373667 CXC64;

```

```

Query Match 64.9%; Score 37; DB 1; Length 423;
Best Local Similarity 85.7%; Pred. No. 19;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 5 YTAPOPO 11
Db 89 YTAPOPE 95

```

```

RESULT 4
STRN_STRGR STANDARD; PRT; 384 AA.
ID STRN_STRGR
AC P09399;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE dtdp-dihydrostreptose-streptidine-6-phosphate
DE dihydrostreptocytosyltransferase (EC 2.4.2.27).
GN STRN.
OS Streptomyces griseus.
OC Bacteria; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomycetes.
OX NCBI_TaxID=1911;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-N2-3-11;
RX MEDLINE=91375432; PubMed=1654502;
RA Mansouri K., Plepersberg W.;
RT "Genetics of streptomycin production in Streptomyces griseus:
RT nucleotide sequence of five genes, strFgH1K, including a phosphatase
RT gene."
RL Mol. Gen. Genet. 228:459-469(1991).
CC -1- FUNKTION: PROBABLY IS A DIHYDROSTREPTOSYL GLYCOSYLTRANSFERASE,
CC INVOLVED IN THE FIRST GLYCOSYLATION STEP CONDENSING STREPTIDINE-
CC 6-PHOSPHATE AND DIHYDROSTREPTOSE (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: dtdp-L-dihydrostreptose + streptidine 6-
CC phosphate = dtdp + 0-1,4-alpha-L-dihydrostreptosyl-streptidine 6-
CC phosphate.
CC -1- PATHWAY: Streptomycin biosynthesis.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

```

```

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
DR EMBL: Y00459; CAA68520.1;
DR PIR: S17778; S17778.
KW Streptomycin biosynthesis; Transferase; Glycosyltransferase.
SQ SEQUENCE 384 AA; 42900 MW; 61BB01FB77222A71 CXC64;

```

```

Query Match 63.2%; Score 36; DB 1; Length 384;
Best Local Similarity 70.0%; Pred. No. 26;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 2 ASSTAPPO 11
Db 181 ASYSTAPPE 190

```

```

RESULT 5
CP24_HUMAN STANDARD; PRT; 514 AA.
ID CP24_HUMAN
AC 007973; Q15807;
DT 01-OCT-1996 (Rel. 34, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Cytochrome P450-CC24, mitochondrial precursor (EC 1.14.-.-) (P450-
DE CC24) (Vitamin D(3) 24-hydroxylase) (1,25-dihydroxyvitamin D(3) 24-
DE hydroxylase) (24-OHASE).
GN CYP24.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=93281615; PubMed=8506296;
RA Chen K.-S., Prah J.M., Deluca H.F.;
RT "Isolation and expression of human 1,25-dihydroxyvitamin D3 24-
RT hydroxylase cDNA."
RL Proc. Natl. Acad. Sci. U.S.A. 90:4543-4547(1993).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21638749; PubMed=11780052;
RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,
RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Baguley C.L.,
RA Bailey J., Barlow K.E., Bates K.N., Beard L.M., Beare D.M.,
RA Beasley O.P., Bird C.P., Blake S.E., Bridgeman A.M., Brown A.J.,
RA Chapman J.C., Clamp M., Butler A.P., Carder C., Carter N.P.,
RA Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R.,
RA Coulson A., Coville G.J., Deadman R., Dhani P.D., Dunn M.,
RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,
RA Graham D.V., Griffiths C., Griffiths M.N.D., Gilliam R., Hall R.E.,
RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.E., Howden P.J.,
RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,
RA Key M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,
RA Levasialho M.H., Leverisha M.A., Lloyd C., Lloyd D.M., Lovell J.A.,
RA Marsh V.L., Martin S.L., McCormack L.J., McLeay K., Murray A.D.,
RA Milne S.A., Mistry D., Moore M.-J.F., Mullikin J.C., Nickerson T.,
RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,
RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramay H.,
RA Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Showkhen R., Sims S.,
RA Skuce C.D., Smith M.L., Soderlund C., Stewart C.A., Sulston J.E.,
RA Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,
RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,
RA Whitehead S.L., Whitaker P., Willey D.L., Williams L., Williams S.A.,
RA Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,
RA Rogers J.,
RT "The DNA sequence and comparative analysis of human chromosome 20."
RL Nature 414:865-871(2001).
RN [3]
RP SEQUENCE OF 1-104 FROM N.A.

```

RX MEDLINE=95359195; PubMed=7632726;
 RA Chen K.-S., Deluca H.F.;
 RT "Cloning of the human I alpha,25-dihydroxyvitamin D-3 24-hydroxylase
 RT gene promoter and identification of two vitamin D-responsive
 RT elements.";
 RL Biochim. Biophys. Acta 1263:1-9(1995).
 RN [4]
 RP SEQUENCE OF 258-514 FROM N.A.
 RX MEDLINE=94091187; PubMed=826631;
 RA Labuda M., Lemieux N., Tily F., Prinster C., Giorleux F.H.;
 RT "Human 25-hydroxyvitamin D 24-hydroxylase cytochrome P450 subunit
 RT maps to a different chromosomal location than that of pseudovitamin
 RT D-deficient rickets.";
 RL J. Bone Miner. Res. 8:1397-1406(1993).
 CC -1- FUNCTION: HAS A ROLE IN MAINTAINING CALCIUM HOMEOSTASIS. CATALYZES
 CC THE NADPH-DEPENDENT 24-HYDROXYLATION OF 25-HYDROXYVITAMIN D(3) IN
 CC THE PRESENCE OF ADRENODOXIN AND NADPH-ADRENODOXIN REDUCTASE.
 CC -1- SUBCELLULAR LOCATION: Mitochondrial.
 CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC
 DR EMBL; L13286; AAA62379.1; -;
 DR EMBL; AL138805; CAB91829.1; -;
 DR EMBL; U60669; AAB03776.1; ALT_SEQ.
 DR EMBL; S67623; AAB29308.1; -;
 DR HSSP; P14779; LJP2.
 DR Genew; HGNC:2602; CYP24.
 DR MIM; 126065; -;
 DR InterPro; IPR001128; Cytochrome_P450.
 DR Pfam; PF00067; P450; 1.
 DR PRINTS; PR00385; P450.
 DR PROSITE; PS00086; CYTOCHROME_P450; 1.
 KW Oxidoreductase; Monooxygenase; Electron transport; Heme;
 KM Mitochondrion; Transit peptide.
 FT TRANSIT 1 35 MITOCHONDRION (BY SIMILARITY).
 FT CHAIN 36 514 CYTOCHROME P450-CC24.
 FT BINDING 462 462 HEME (POTENTIAL).
 FT CONFLICT 68 68 G -> A (IN REF. 1).
 FT CONFLICT 124 125 AY -> V (IN REF. 1).
 FT CONFLICT 270 270 D -> G (IN REF. 1).
 FT CONFLICT 365 365 V -> R (IN REF. 1).
 FT CONFLICT 368 368 A -> E (IN REF. 1).
 FT CONFLICT 390 390 S -> G (IN REF. 1).
 FT CONFLICT 511 511 F -> S (IN REF. 4).
 SQ SEQUENCE 514 AA; 58875 MW; 8662F63771981195 CRC64;
 Query Match 63.2%; Score 36; DB 1; Length 514;
 Best Local Similarity 50.0%; Pred. No. 35;
 Matches 5; Conservative 5; Mismatches 0; Indels 0; Gaps 0;
 QY 2 ASSTYRPOP 11
 Db 30 STAYTSPQPR 39
 RESULT 6
 ID TRA_BPMU STANDARD; PRT; 663 AA.
 AC P07636; P06021;
 DT 01-APR-1988 (Rel. 07, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Transposase.
 GN A OR 3.
 OS Bacteriophage Mu.
 OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Myoviridae;

OC Mu-like viruses.
 OX NCBI_TaxID=10677;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=86067968; PubMed=2999776;
 RA Harshay R.M., Getzoff E.D., Baldwin D.L., Miller J.L., Chaconas G.;
 RT "Primary structure of phage mu transposase: homology to mu
 RT repressor.";
 RL Proc. Natl. Acad. Sci. U.S.A. 82:7676-7680(1985).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Priess H., Brauer B., Schmidt C., Kamp D.;
 RT "Sequence of the left end of Mu.";
 RL (in) Symonds N., Roussaint A., Van de Putte P., Howe M.M. (eds.);
 RL phage Mu, pp.277-296, Cold Spring Harbor Laboratory Press,
 RL New York (1987).
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Morgan G., Hattall G., Hendrix R.;
 RT "Genome of bacteriophage Mu and comparison with the Haemophilus
 RT influenzae Mu-like prophage Flumu.";
 RL Submitted (Aug-1998) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE OF 1-88 FROM N.A.
 RX MEDLINE=83012203; PubMed=6214696;
 RA Priess H., Kamp D., Kahmann R., Brauer B., Delius H.;
 RT "Nucleotide sequence of the immunity region of bacteriophage Mu";
 RL Mol. Gen. Genet. 186:315-321(1982).
 RN [5]
 RP SEQUENCE OF 1-84 FROM N.A.
 RX MEDLINE=83218562; PubMed=6222246;
 RA Toussaint A., Faellen M., Desmet L., Allet B.;
 RT "The products of gene A of the related phages Mu and D108 differ in
 RT their specificities.";
 RL Mol. Gen. Genet. 190:70-79(1983).
 RN [6]
 RP STRUCTURE BY NMR OF 1-76.
 RX MEDLINE=95187707; PubMed=7881904;
 RA Clubb R.T., Omichinski J.G., Savillanti H., Mizuuchi K.,
 RA Gronenborn A.M., Clore G.M.;
 RT "A novel class of winged helix-turn-helix protein: the DNA-binding
 RT domain of Mu transposase.";
 RL Structure 2:1041-1048(1994).
 RN [7]
 RP STRUCTURE BY NMR OF 76-174.
 RX MEDLINE=98070329; PubMed=9405381;
 RA Schumacher S., Clubb R.T., Cai M., Mizuuchi K., Clore G.M.,
 RA Gronenborn A.M.;
 RT "Solution structure of the Mu end DNA-binding Ibeta subdomain of
 RT phage Mu transposase: modular DNA recognition by two tethered
 RT domains.";
 RL EMBO J. 16:7532-7541(1997).
 RN [8]
 RP STRUCTURE BY NMR OF 173-247.
 RX MEDLINE=98035037; PubMed=9367742;
 RA Clubb R.T., Schumacher S., Mizuuchi K., Gronenborn A.M., Clore G.M.;
 RT "Solution structure of the I gamma subdomain of the Mu end
 RT DNA-binding domain of phage Mu transposase.";
 RL J. Mol. Biol. 273:19-25(1997).
 RN [9]
 RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS) OF 248-574.
 RX MEDLINE=95354202; PubMed=7628012;
 RA Rice P., Mizuuchi K.;
 RT "Structure of the bacteriophage Mu transposase core: a common
 RT structural motif for DNA transposition and retroviral integration.";
 RL Cell 82:209-220(1995).
 CC -1- FUNCTION: THIS TRANSPOSASE IS ESSENTIAL FOR INTEGRATION,
 CC REPLICATION-TRANSPOSITION, AND EXCISION OF MU DNA.
 CC -1- MISCELLANEOUS: MU CAN TRANSPOSE ITS DNA INTO MULTIPLE SITES IN
 CC MANY BACTERIAL GENOMES AND MEDIATE A VARIETY OF DNA
 CC REARRANGEMENTS. TRANSPOSITION REQUIRES BOTH TRANSPOSASE (ENCODED
 CC BY GENE A) AND TRANSPOSITION ENHANCER (ENCODED BY GENE B).
 CC -1- MISCELLANEOUS: UNLIKE OTHER TRANSPOSONS MU HAS DISSIMILAR

CC SEQUENCES AT ITS LEFT AND RIGHT ENDS. TRANSPOSASE APPARENTLY BINDS
 CC 3 SPECIFIC BLOCKS OF SEQUENCES AT EACH END OF MU DNA.
 CC -1- MISCELLANEOUS: THE A GENE IS REGULATED BY THE REPRESSOR C, WHICH
 CC BINDS TO AN OPERATOR SEQUENCE & TURNS OFF TRANSCRIPTION. REPRESSOR
 CC C CAN, AT HIGH CONCENTRATIONS, OCCUPY ALMOST THE EXACT SAME SITES
 CC ON MU ENDS AS THE TRANSPOSASE, AND TRANSPOSASE CAN BIND TO
 CC FRAGMENTS CONTAINING THE MU OPERATOR SEQUENCE.
 CC -1- SIMILARITY: STRONG, TO H. INFLUENZAE H11478.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: M1195; AAA2369.1; -
 CC EMBL: M64097; AAA2379.1; -
 CC EMBL: AF083977; AAF01083.1; -
 CC EMBL: V01464; CAA24713.1; -
 CC EMBL: V00868; CAA24236.1; -
 CC PIR: A24746; TOBPU.
 CC DR PIR: 1TNS; 14-FEB-95.
 CC DR PDB: 1TNT; 14-FEB-95.
 CC DR PDB: 1BCM; 15-OCT-95.
 CC DR PDB: 1BCO; 15-OCT-95.
 CC DR PDB: 2EZH; 03-DEC-97.
 CC DR PDB: 2E2I; 03-DEC-97.
 CC DR PDB: 2E2K; 14-JAN-98.
 CC DR PDB: 2E2L; 14-JAN-98.
 CC DR InterPro: IPR003314; Mu_DNA_bind.
 CC DR InterPro: IPR004189; Mu_transposase.
 CC DR Pfam: PF02316; Mu_DNA_bind.1;
 CC DR Pfam: PF02914; Mu_transposase.1;
 CC KW Transposition; Transposable element; DNA-binding; DNA excision;
 CC DNA integration; DNA recombination; 3D-structure.
 CC FT DNA_BIND 35 55 H-T-H MOTIF (POTENTIAL).
 CC FT DNA_BIND 390 409 H-T-H MOTIF (POTENTIAL).
 CC FT CONFLICT 66 66 G -> R (IN REF. 5).
 CC FT CONFLICT 408 408 P -> S (IN REF. 2).
 CC SQ SEQUENCE 663 AA; 75003 MW; B862CFDCEFC082E3 CRC64;
 CC
 CC Query Match 63.2%; Score 36; DB 1; Length 663;
 CC Best Local Similarity 60.0%; Pred. No. 46;
 CC Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
 CC
 CC QY 2 ASSTAPQPO 11
 CC DB 411 AGAYTGPMPQ 420
 CC
 CC RESULT 7
 CC CASB_CAMDR STANDARD; PRT; 232 AA.
 CC AC Q9TVDO;
 CC DT 16-OCT-2001 (Rel. 40, Created)
 CC DT 16-OCT-2001 (Rel. 40, Last sequence update)
 CC DT 16-OCT-2001 (Rel. 40, Last annotation update)
 CC DE Beta casein precursor.
 CC GN CSN2.
 CC OS Camelus dromedarius (Dromedary) (Arabian camel).
 CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC OC Mammalia; Eutheria; Cetartiodactyla; Tylopoda; Camelidae; Camelus.
 CC NCBI_TaxID=9838;
 CC RN [1]
 CC RP SEQUENCE FROM N.A.
 CC RC STRAIN=Somali; TISSUE=Mammary gland;
 CC RX MEDLINE=98291310; PubMed=9627840;
 CC RA Kappellet S., Farah Z., Puhon Z.,
 CC RA "Sequence analysis of Camelus dromedarius milk caseins".
 CC RA J. Dairy Res. 65:209-222(1998).
 CC RL -1- FUNCTION: IMPORTANT ROLE IN DETERMINATION OF THE SURFACE

CC PROPERTIES OF THE CASEIN MICELLES (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Extracellular.
 CC -1- TISSUE SPECIFICITY: MAMMARY GLAND; MILK.
 CC -1- SIMILARITY: BELONGS TO THE BETA-CASEIN FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: AJ012630; CAA10079.1; -
 CC DR InterPro: IPR001588; Casein.
 CC DR Pfam: PF00363; caseins.1.
 CC KW PROSITE: PS00306; CASEIN_ALPHA_BETA; FALSE_NEG.
 CC DR MILK; Phosphorylation; Glycoprotein; Signal.
 CC FT SIGNAL 1 15 BY SIMILARITY.
 CC FT CHAIN 16 232 BETA CASEIN.
 CC FT MOD_RES 30 30 PHOSPHORYLATION (BY SIMILARITY).
 CC FT MOD_RES 32 32 PHOSPHORYLATION (BY SIMILARITY).
 CC FT MOD_RES 33 33 PHOSPHORYLATION (BY SIMILARITY).
 CC FT MOD_RES 34 34 PHOSPHORYLATION (BY SIMILARITY).
 CC SQ SEQUENCE 232 AA; 26218 MW; A0F9F41D2EA7C518 CRC64;
 CC
 CC Query Match 61.4%; Score 35; DB 1; Length 232;
 CC Best Local Similarity 63.6%; Pred. No. 23;
 CC Matches 7; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 CC
 CC QY 1 QASSTAPQPO 11
 CC DB 62 QDKITFPQPO 72
 CC
 CC RESULT 8
 CC IF2M_SCHPO STANDARD; PRT; 686 AA.
 CC AC IF2M_SCHPO
 CC ID 059683;
 CC DT 15-DEC-1998 (Rel. 37, Created)
 CC DT 15-DEC-1998 (Rel. 37, Last sequence update)
 CC DT 15-JUN-2002 (Rel. 41, Last annotation update)
 CC DE Translation initiation factor IF-2, mitochondrial precursor (IF-2Mt)
 CC DE (IF-2Mt).
 CC GN SPBC2F6.01 OR SPBC1271.15C.
 CC OS Schizosaccharomyces pombe (Fission yeast).
 CC OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 CC OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 CC OC Schizosaccharomycetes.
 CC NCBI_TaxID=4896;
 CC RN [1]
 CC RP SEQUENCE FROM N.A.
 CC RC STRAIN=972;
 CC RX MEDLINE=21846401; PubMed=11859360;
 CC RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
 CC RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
 CC RA Brooks K., Brown D., Brown S., Chillingworth T., Church C.M.,
 CC RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
 CC RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
 CC RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagsels K.,
 CC RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
 CC RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
 CC RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinovitch E.,
 CC RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
 CC RA Skelton J., Simmonds M., Squares R., Stevens K.,
 CC RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
 CC RA Woodward J., Volckaert G., Aert R., Robben S., Gymnopoulos B.,
 CC RA Wellens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
 CC RA Gabel C., Fuchs M., Fritze C., Holzer E., Moesti D., Hilbert H.,
 CC RA Borzym K., Langer I., Beck A., Lehnach H., Reinhardt R., Pohl T.M.,
 CC RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
 CC RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaire V., Mottler S.,
 CC RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,

RT filament): a molecular leaf spring.";
RN
RL J. Cell Biol. 111:11089-1105(1990).
[2]
RN SEQUENCE FROM N.A.
RX MEDLINE=96311563; PubMed=8733135;
RA Chen E.Y., Zollo M., Mazzarella R.A., Ciocodico A., Chen C.-N.,
RA Zuo L., Heiner C., Burrough F.W., Ripetto M., Schlessinger D.,
RA D'Urso M.;
RT "Long-range sequence analysis in Xq28: thirteen known and six
RT candidate genes in 219.4 kb of high GC DNA between the RCP/SCP and
RT 66PD loci.";
RN Hum. Mol. Genet. 5:659-668(1996).
[3]
RN SEQUENCE OF 1658-1772 FROM N.A.
RX MEDLINE=93357748; PubMed=7689010;
RA Maestrini E., Patrosso C., Mancini M., Rivella S., Rocchi M.,
RA Ripetto M., Valla A., Fratini A., Zoppe M., Vezoni P.,
RA Tonolo D.;
RT "Mapping of two genes encoding isoforms of the actin binding protein
RL ABP-280, a dystrophin like protein, to Xq28 and to chromosome 7.";
RN Hum. Mol. Genet. 2:761-766(1993).
[4]
RN REVIEW.
RX MEDLINE=21234905; PubMed=11336782;
RA van der Flier A., Sonnenberg A.;
RT "Structural and functional aspects of filamins.";
RL Biochim. Biophys. Acta 1538:99-117(2001).
[5]
RN REVIEW.
RX MEDLINE=21146932; PubMed=11252955;
RA Stossel T.P., Condeelis J., Cooley L., Hartwig J.H., Noegel A.,
RT Schleicher M., Shapiro S.S.;
RL "Filamins as integrators of cell mechanics and signalling.";
RN Nat. Rev. Mol. Cell Biol. 2:338-145(2001).
[6]
RN VARIANTS PH PHE-656 AND THR-1764.
RX MEDLINE=2123782; PubMed=11533987;
RA Sheen V.L., Dixon P.H., Fox J.W., Hong S.E., Kinton L., Sisodia S.M.,
RA Duncan J.S., Dubau F., Scheffer I.E., Schachter S.C., Wilner A.,
RA Henchy R., Cirio P., Kamuro K., Dimario F., Berg M., Kuzniack R.,
RA Cole A.J., Bromfield E., Biber M., Schomer D., Wheelless J., Silver R.,
RA Mochida G.H., Berkovic S.F., Andermann F., Andermann E., Dobyns W.B.,
RA Wood N.W., Walsh C.A.;
RT "Mutations in the X-linked filamin 1 gene cause periventricular
RN nodular heterotopia in males as well as in females.";
RL Hum. Mol. Genet. 10:1775-1783(2001).
[7]
RN FUNCTION: Promotes orthogonal branching of actin filaments and
RL links actin filaments to membrane glycoproteins. Anchors various
CC transmembrane proteins to the actin cytoskeleton and serves as a
CC scaffold for a wide range of cytoplasmic signaling proteins.
CC
CC -1 SUBUNIT: Homodimer. Interacts with cWSP. Interacts with various
CC other binding partners in addition to filamentous actin.
CC
CC -1 SUBCELLULAR LOCATION: PERIPHERAL CYTOPLASM.
CC
CC -1 TISSUE SPECIFICITY: UBIQUITOUS.
CC
CC -1 PIM: PHOSPHORYLATION EXTENT CHANGES IN RESPONSE TO CELL
CC
CC ACTIVATION.
CC
CC -1 PIM: The N-terminus is blocked.
CC
CC -1 DISEASE: Defects in FLNA are the cause of periventricular
CC heterotopia (PH) also called nodular heterotopia, bilateral
CC periventricular (NBP or BPN). PH is an X-linked developmental
CC dominant disorder in which many neurons fail to migrate into the
CC cerebral cortex. They remain as nodules lining the ventricular
CC surface. In heterozygous females these neurons presumably
CC represent those cells that, after X-chromosome inactivation,
CC contain the active X chromosome with the filamin mutation. Most
CC hemizygous affected males die early during embryogenesis, whereas
CC heterozygous females have normal intelligence but suffer from
CC seizures and various manifestations outside the central nervous
CC system, especially related to the vascular system. This implies
CC that essential embryonic cell migration can only occur in FLNA-
CC expressing cells.
CC
CC -1 SIMILARITY: THE ACTIN-BINDING DOMAIN IS OF A TYPE FOUND IN MANY
CC ACTIN-BINDING PROTEINS (SUCH AS ACTININ, DYSTROPHIN, FIMBRIN,

ABP-120, ABP-180, OR BETA-FODRIN).
 -1- SIMILARITY: CONTAINS 2 CALPONIN-HOMOLOGY (CH) DOMAINS.
 -1- SIMILARITY: CONTAINS 2 FILAMIN REPEATS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: X53416; CAA37495.1; -
 DR EMBL: L44140; AAA92844.1; -
 DR EMBL: X70082; CAA45687.1; -
 DR PIR: A37098; CAA49690.1; -
 DR HSSP: P13466; 1KSR.
 DR Genbank: HGNC:3754; FLNA.
 DR MIM: 300017; -
 DR MIM: 300049; -
 DR InterPro: IPR001589; Actbind_actuin.
 DR InterPro: IPR001715; Calponin-like.
 DR Pfam: PF00307; CH; 2.
 DR SMART: SM00033; CH; 2.
 DR PROSITE: PS00019; ACTININ_1; 1.
 DR PROSITE: PS00020; ACTININ_2; 1.
 DR PROSITE: PS50194; FILAMIN_REPEAT; 24.
 KW Actin-binding; Phosphorylation; Repeat; Polymorphism;
 FT DOMAIN 1 274
 FT DOMAIN 1 149
 FT DOMAIN 1 166
 FT REPEAT 276 266
 FT REPEAT 374 474
 FT REPEAT 475 570
 FT REPEAT 571 663
 FT REPEAT 667 763
 FT REPEAT 764 866
 FT REPEAT 867 965
 FT REPEAT 966 1061
 FT REPEAT 1062 1154
 FT REPEAT 1155 1249
 FT REPEAT 1250 1349
 FT REPEAT 1350 1442
 FT REPEAT 1443 1539
 FT REPEAT 1540 1636
 FT REPEAT 1649 1740
 FT DOMAIN 1741 1778
 FT REPEAT 1779 1860
 FT REPEAT 1861 1950
 FT REPEAT 1951 2039
 FT REPEAT 2042 2131
 FT REPEAT 2132 2230
 FT REPEAT 2233 2325
 FT REPEAT 2327 2420
 FT REPEAT 2424 2516
 FT REPEAT 2517 2551
 FT REPEAT 2552 2646
 FT DOMAIN 2647 2662
 FT SITE 1761 1762
 FT VARIANT 320 320
 FT VARIANT 370 370
 FT VARIANT 552 552
 FT VARIANT 656 656
 FT VARIANT 1764 1764

ACTIN-BINDING (HEAD).
 CH 1.
 CH 2.
 FILAMIN 1.
 FILAMIN 2.
 FILAMIN 3.
 FILAMIN 4.
 FILAMIN 5.
 FILAMIN 6.
 FILAMIN 7.
 FILAMIN 8.
 FILAMIN 9.
 FILAMIN 10.
 FILAMIN 11.
 FILAMIN 12.
 FILAMIN 13.
 FILAMIN 14.
 FILAMIN 15.
 HINGE 1.
 FILAMIN 16.
 FILAMIN 17.
 FILAMIN 18.
 FILAMIN 19.
 FILAMIN 20.
 FILAMIN 21.
 FILAMIN 22.
 FILAMIN 23.
 HINGE 2.
 FILAMIN 24.
 SELF-ASSOCIATION SITE, TAIL.
 CLEAVAGE (BY CALPAIN).
 V->A (IN DBSNP:1064816).
 F->L (IN DBSNP:1064817).
 F->L (IN DBSNP:1064817).
 V->A (IN DBSNP:730319).
 F->F (IN PH).
 L->F (IN PH).
 F->F (IN PH).
 A->T (IN PH).

FT FT
 FT CONFLICT 1772 1772 /FTId-VAR.012835.
 FT CONFLICT 2634 2634 A->G (IN REF. 3).
 SO SEQUENCE 2647 AA; 280759 MW; 6C1A07041DPA3042 CRC64;
 Query Match 61.4%; Score 35; DB 1; Length 2647;
 Best Local Similarity 100.0%; Pred. No. 3e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 5 YTAAP 10
 DB 1712 YTAAP 1717

RESULT 10
 ID PKAA_STRCO STANDARD; PRT; 543 AA.
 AC P54739;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DE 15-JUN-2002 (Rel. 41, Last annotation update)
 GN Serine/threonine protein kinase Pkax (EC 2.7.1.-).
 OS PKAA OR SCO2974 OR SCE50.02C.
 OC Streptomyces coelicolor.
 OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
 OC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.
 ON NCBI_TaxID:1902;
 RP [1]
 RC SEQUENCE FROM N.A.
 RX STRAIN-A3(2);
 RX MEDLINE=9518913; PubMed=7883195;
 RA Urahe H., Ogawara H.;
 RT "Cloning, sequencing and expression of serine/threonine kinase-
 RL gene 153:99-104(1995)."
 RN [2]
 RP SEQUENCE FROM N.A.
 RX STRAIN-A3(2) / M145;
 RX MEDLINE=21996410; PubMed=12000953;
 RA Bentley S.D., Chater K.F., Cerdeno-Parraga A.-M., Challis G.L.,
 RA Thompson N.R., James K.D., Harris D.E., Quail M.A., Chen C.W., Collins M.,
 RA Harper D., Bateman A., Brown S., Chandra G., Hornsby T., Howarth S.,
 RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
 RA Huang C.-H., Kleser T., Larke L., Murphy L., Oliver K., O'Neill S.,
 RA Rabinowitz E., Rajandream M.A., Rutherford K., Rutter S.,
 RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
 RA Warren T., Wietzorek A., Woodward J., Barrell B.G., Parkhill J.,
 RA Hopwood D.A.;
 RT "Complete genome sequence of the model actinomycete Streptomyces
 RL coelicolor A3(2)."
 RT Nature 417:141-147(2002).
 RL -1- PWM: AUTOPHOSPHORYLATED MAINLY AT THR AND SLIGHTLY AT SER.
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: D86821; BA013168.1; -
 DR EMBL: AL163672; CAB87324.1; -
 DR HSSP: O63450; 1A06.
 DR PhosSite: P54739; -
 DR InterPro: IPR000719; Euk_Pkinase.
 DR InterPro: IPR002290; Ser_thr_kinase.
 DR Pfam: PF00069; Pkinase; 1.
 DR ProDom: PD000001; Euk_pkinase; 1.
 DR SMART: SM00220; S_TKc; 1.
 DR PROSITE: PS00107; PROTEIN_KINASE_ATP; FALSE_NEG.
 DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.

DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
 RP Transferase; Serine/threonine-protein kinase; ATP-binding;
 KW Phosphorylation; Complete proteome.
 FT DOMAIN 8 276 PROTEIN_KINASE.
 FT NP_BIND 14 22 ATP (BY SIMILARITY).
 FT BINDING 48 48 ATP (BY SIMILARITY).
 FT ACT_SITE 142 142 BY SIMILARITY.
 FT DOMAIN 301 481 GLN/PRO-RICH
 FT SEQUENCE 543 AA; 58181 MM; 0E1965520FA0C200 CRC64;
 Query Match 59.6%; Score 34; DB 1; Length 543;
 Best Local Similarity 54.5%; Pred. No. 86;
 Matches 6; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 QASSTAPQPO 11
 DB 437 QPQRYATPQPO 447

RESULT 11
 VE4_HPV54 STANDARD; PRT; 134 AA.
 ID VE4_HPV54
 AC 081022;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE Probable E4 protein.
 GN E4.
 OS Human papillomavirus type 54.
 OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
 OC Papillomavirus.
 OC NCBI_TaxID=37113;
 RX [1]
 RP SEQUENCE FROM N.A.

RA Delius H.;
 RL Submitted (Oct-1995) to the EMBL/GenBank/DDBJ databases.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 CC EMBL; U37488; AA79191.1;
 DR InterPro: IPR003861; Papilloma_E4.
 DR Pfam: PF02711; Pap_E4; 1.
 KW Early protein.
 SQ SEQUENCE 134 AA; 15373 MM; B157979B6035EE2A CRC64;

Query Match 57.9%; Score 33; DB 1; Length 134;
 Best Local Similarity 66.7%; Pred. No. 30;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 QASSTAPQ 9
 DB 96 QTSNHTAPQ 104

RESULT 12
 CPQ3_CANPG STANDARD; PRT; 175 AA.
 ID CPQ3_CANPG
 AC P81584;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Cuticle protein CP1876 (CPCL1876).
 DE Cancer pagurus (Rock crab).
 OS Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Crustacea;
 OC Malacostraca; Eumalacostraca; Decapoda; Pleocyemata;
 OC Brachyura; Eubrachyura; Cancroidea; Cancridae; Cancer.
 OC NCBI_TaxID=6755;

RN [1]
 RP SEQUENCE.
 RC TISSUE-Cuticle;
 RX MEDLINE=99354472; PubMed=10425740;
 RA Andersen S.O.;
 RT "Exoskeletal proteins from the crab, Cancer pagurus.";
 RL Comp. Biochem. Physiol. 123A:203-211(1999).
 CC -1- TISSUE SPECIFICITY: CALCIFIED SHELL.
 CC -1- MASS SPECTROMETRY: MW=18759.8; METHOD=MALDI.
 KW Structural protein; Cuticle.
 FT SEQUENCE 175 AA; 18765 MM; 57C1C83D75CB132E CRC64;
 Query Match 57.9%; Score 33; DB 1; Length 175;
 Best Local Similarity 60.0%; Pred. No. 40;
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 ASSYAPQPO 11
 DB 43 ADVYMPRPQ 52

RESULT 13
 FL3L_HUMAN STANDARD; PRT; 235 AA.
 ID FL3L_HUMAN
 AC P49771;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE SL cytokine precursor (Fms-related tyrosine kinase 3 ligand) (Flt3
 DE ligand) (Flt3L).
 GN FLT3LG.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OC NCBI_TaxID=9606;
 RX [1]
 RP SEQUENCE FROM N.A.

RA Hannum C., Culpepper J., Campbell D., McClanahan T., Zurawski S.,
 RA Hannum C., Culpepper J., Campbell D., McClanahan T., Zurawski S.,
 RA Bada G., Martina N., Peterson D., Menon S., Shanaleit A.,
 RA Muench A., Kelnier G., Nankawa R., Rennick D., Roncarolo M.G.,
 RA Zlotnick A., Rosnet O., Dubreuil P., Barnbaum D., Lee F.;
 RA "Ligand for FLT3/FLT3 receptor tyrosine kinase regulates growth of
 RA haematopoietic stem cells and is encoded by variant RNAs.";
 RL Nature 368:643-648(1994).
 RL
 RL
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94155428; PubMed=8145851;
 RA Lyman S.D., James L., Johnson L., Brasel K., de Vries P.,
 RA Escobar S.S., Downey H., Splet R.R., Beckmann M.P., McKenna H.J.;
 RA "Cloning of the human homologue of the murine flt3 ligand: a growth
 RA factor for early hematopoietic progenitor cells.";
 RL Blood 83:2795-2801(1994).
 RL
 RP SEQUENCE FROM N.A. AND ALTERNATIVE SPLICING.
 RX MEDLINE=96032581; PubMed=7566977;
 RA Lyman S.D., Stocking K., Davison B., Fletcher F., Johnson L.,
 RA Escobar S.S.;
 RL "Structural analysis of human and murine flt3 ligand genomic loci.";
 RL Oncogene 11:1165-1172(1995).
 RL [4]
 RN X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS).
 RP MEDLINE=20343011; PubMed=10881197;
 RX Savvides S.N., Boone T., Karplus P.A.;
 RA "Flt3 ligand structure and unexpected commonalities of helical
 RA bundles and cystine knots.";
 RL Nat. Struct. Biol. 7:486-491(2000).
 CC -1- FUNCTION: STIMULATES THE PROLIFERATION OF EARLY HEMATOPOIETIC
 CC CELLS. SYNERGIZES WELL WITH A NUMBER OF OTHER COLONY STIMULATING
 CC FACTORS AND INTERLEUKINS.
 CC -1- SUBUNIT: Homodimer (isoform 2).
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein (isoform 1);

CC secreted (isoform 2).

CC -1- ALTERNATIVE PRODUCTS: 2 isoforms; a 1/membrane-bound (shown here)

CC and 2/soluble; are produced by alternative splicing.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: U04806; AAA17999.1; -
DR EMBL: U03858; AAA19825.1; -
DR EMBL: U29874; AAA90949.1; -
DR PDB: 1ETE; 09-JUN-00.
DR Genbank: HGNC:3766; FLT3LG.
DR MIM: 600007; -
DR InterPro: IPR004213; FLT3_L19.
DR Pfam: PF02947; flt3_L19; 1.
KW Cytokine; Glycoprotein; Transmembrane; Alternative splicing; Signal;
KW 3D-structure.

FT SIGNAL 1 26
FT CHAIN 27 235
FT DOMAIN 27 184
FT TRANSMEM 185 205
FT DOMAIN 206 235
FT DISULFID 30 111
FT DISULFID 70 153
FT DISULFID 119 158
FT CARBOHD 126 126
FT CARBOHD 149 149
FT VARSPLIC 161 178
FT VARSPLIC 179 235
FT VARSPLIC 72 235
FT VARSPLIC 235 AA; 26416 MW; 738958P6934CECF CRC64;
SQ SEQUENCE

Query Match
Best Local Similarity 57.9%; Score 33; DB 1; Length 235;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
QY 1 QASSYTAPOP 10
Db 175 EATVAPAP 184

RESULT 14
OMPY_CHLPN
ID OMPEY_CHLPN STANDARD; PRT; 262 AA.
AC 0926M5;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE Putative outer membrane protein Cpn1034/CP0818/CPJ1034 precursor.
GN Cpn1034 OR CP0818 OR CPJ1034.
OS Chlamydia pneumoniae (Chlamydia pneumoniae).
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=83558;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-CM1029;
RX MEDLINE=99206606; PubMed=10192388;
RA Kallmar S., Mitchell W., Marathe R., Lammel C., Fan J., Hyman R.W.,
RT Olinger L., Grimwood J., Davis R.W., Stephens R.S.,
RL "Comparative genomes of Chlamydia pneumoniae and C. trachomatis.",
J. Nat. Genet. 21:385-389(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-AR39;
RX MEDLINE=20150255; PubMed=10684935;
RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,

RA White O., Hickey E.K., Peterson J., Utterback T., Berry K., Bass S.,
RA Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,
RA Glinn M., Nelson W., Deboy R., Kolony J., McClarty G., Salzberg S.L.,
RT "Genome sequences of Chlamydia trachomatis Mopn and Chlamydia
RT pneumoniae AR39.",
RN Nucleic Acids Res. 28:1397-1406(2000).
RP [3]
RC SEQUENCE FROM N.A.
RX STRAIN-J138;
RX MEDLINE=2030349; PubMed=10871362;
RA Shiba T., Ishii K., Hattori M., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,
RT "Comparison of whole genome sequences of Chlamydia pneumoniae J138
RT from Japan and CW1029 from USA.",
RN Nucleic Acids Res. 28:2311-2314(2000).
CC -1- SUBCELLULAR LOCATION: Outer membrane (Potential).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: AE001684; AAD19171.1; -
DR EMBL: AE002240; AAR38613.1; -
DR EMBL: AP002548; BAA99241.1; -
DR TIGR: CP0818; -
KW Hypothetical protein; Outer membrane; Signal; Complete proteome.
FT CHAIN 1 17
FT SIGNAL 18 262
FT SIGNAL 18 262
SQ SEQUENCE 262 AA; 30000 MW; 23D30M12P6C5A893 CRC64;
CPN1034/CP0818/CPJ1034.

Query Match
Best Local Similarity 57.9%; Score 33; DB 1; Length 262;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 6 TAPOPQ 11
Db 148 TAPOPQ 153

RESULT 15
GDA4_WHEAT
ID GDA4_WHEAT STANDARD; PRT; 297 AA.
AC P04724;
DT 13-AUG-1987 (Rel. 05, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Alpha/beta-gliadin A-IV precursor (Prolamin).
OS Triticum aestivum (Wheat).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliopsida; Liliopsida; Poales; Poaceae; Poaceae;
OC Triticaceae; Triticum.
OX NCBI_TaxID=4565;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85234522; PubMed=2989281;
RA Ohta T.W., Cheesbrough V., Reeves C.D.,
RT "Evolution and heterogeneity of the alpha-/beta-type and gamma-type
RT gliadin DNA sequences.",
RL J. Biol. Chem. 260:8203-8213(1985).
CC -1- FUNCTION: GLIADIN IS THE MAJOR SEED STORAGE PROTEIN IN WHEAT.
CC -1- MISCELLANEOUS: THE ALPHA/BETA-GLIADIN CAN BE DIVIDED INTO 5
CC HOMOLOGY CLASSES. SEQUENCE DIVERGENCE BETWEEN THE CLASSES IS DUE
CC TO SINGLE BASE SUBSTITUTIONS AND TO DUPLICATIONS OR DELETIONS
CC WITHIN OR NEAR DIRECT REPEATS. THERE ARE MORE THAN A 100 COPIES OF
CC THE GENE FOR ALPHA/BETA-GLIADIN PER HAPLOID GENOME.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; M11075; AAA34282.1; -
 DR PIR; D22364; D22364.
 DR InterPro; IPR003612; AAI.
 DR InterPro; IPR001954; Gila_glutenln.
 DR InterPro; IPR001768; Try/amyln_inhbr.
 DR Pfam; PF00234; clyp_alpha_amy1; 1.
 DR PRINTS; PR00208; GLIADGLUTEN.
 DR SMART; SM00499; AAI; 1.
 KW Seed storage protein; Repeat; Signal; Multigene family.
 FT SIGNAL 1 20
 FT CHAIN 21 297 ALPHA/BETA-GLIADIN A-IV.
 SQ SEQUENCE 297 AA; 34239 MW; 0025ED289AE9588B CRC64;

Query Match 57.9%; Score 33; DB 1; Length 297;
 Best local Similarity 75.0%; Pred. No. 70;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 4 STAPPOQ 11
 II IIII
 DB 106 STPOPOPOQ 113

Search completed: January 13, 2003, 09:54:07
 Job time : 13.3333 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 13, 2003, 09:53:20 ; Search time 10.1538 Seconds
(without alignments)
104.146 Million cell updates/sec

Title: US-09-554-941-2
Perfect score: 57
Sequence: 1 QASSTAPQPQ 11

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues
Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: PIR:73:*
2: PIR:1:*
3: PIR:3:*
4: PIR:4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	52	91.2	728	1 A38084	galactose oxidase
2	40	70.2	1416	2 T20823	hypothetical prote
3	39	68.4	224	2 B75388	NADH dehydrogenase
4	37	64.9	90	2 G75509	hypothetical prote
5	37	64.9	112	2 A43605	potassium efflux s
6	37	64.9	583	2 G87719	protein R119.7 (lm
7	36	63.2	139	2 S31941	penicillin-binding
8	36	63.2	384	2 S17778	strH protein - str
9	36	63.2	482	2 AE3381	ATP-dependent RNA
10	36	63.2	513	1 A47436	1,25-dihydroxyvita
11	36	63.2	546	2 A87372	xylian 1,4-beta-xy
12	36	63.2	662	1 T08PU	transposase - phag
13	36	63.2	681	2 T23455	hypothetical prote
14	36	63.2	1082	2 T42204	chromatin structur
15	35	61.4	483	2 B82722	hypothetical prote
16	35	61.4	686	2 T33951	translation initia
17	35	61.4	820	2 T46412	ubiquitin-protein
18	35	61.4	874	2 T34922	phosphoenolpyruvat
19	35	61.4	875	1 A57080	cell surface antis
20	35	61.4	1124	2 T30340	dsRNA adenosine de
21	35	61.4	2129	2 T27431	hypothetical prote
22	35	61.4	2647	2 T08561	gelation factor AB
23	34	59.6	151	2 T08561	hypothetical prote
24	34	59.6	336	2 E84594	AP2 domain transcr
25	34	59.6	373	2 S72368	single-stranded DN
26	34	59.6	440	2 G84631	hypothetical prote
27	34	59.6	453	2 T17823	purple acid phosph
28	34	59.6	466	1 T04599	hypothetical prote
29	34	59.6	466	1 T04599	acid phosphatase (

30	34	59.6	487	2 H82298	D-alanyl-D-alanine
31	34	59.6	502	2 T08776	hypothetical prote
32	34	59.6	525	2 A75570	2-oxo acid dehydro
33	34	59.6	543	1 A75570	protein kinase (EC
34	34	59.6	616	2 E84424	proline auxin tra
35	34	59.6	640	2 G96733	auxin transport pr
36	34	59.6	706	2 G71004	hypothetical prote
37	34	59.6	818	1 J43397	peroxinectin precu
38	34	59.6	856	2 T13159	E1B-55kDa-associat
39	34	59.6	1012	2 T13159	RAE-28 - mouse
40	34	59.6	1012	2 T43162	vitellinogenin - gyp
41	33	57.9	64	2 AD3228	isopentenyladenosine
42	33	57.9	182	2 G70687	hypothetical prote
43	33	57.9	205	2 A43407	hypothetical cyto
44	33	57.9	235	2 I38440	fl3 ligand - huma
45	33	57.9	251	2 T14456	MADS box protein h

ALIGNMENTS

RESULT 1
A38084
galactose oxidase (EC 1.1.3.9) precursor [validated] - fungus (Cladobotryum dendroides)
C:Species: Cladobotryum dendroides
C:Date: 20-Oct-2000 #sequence_revision 20-Oct-2000 #text_change 20-Oct-2000
R:McPherson, M.J.; Ogel, Z.B.; Stevens, C.; Yadav, K.D.S.; Keen, J.N.; Knowles, P.F.
J. Biol. Chem. 267, 8146-8152, 1992
A:Title: Galactose oxidase of Dactylium dendroides. Gene cloning and sequence analysis.
A:Reference number: A38084; MUID:92235025; PMID:1569070
A:Accession: A38084
A:Molecule type: DNA
A:Residues: 1-728 <MC>
A:Cross-references: GB:M86819
A:Note: it is uncertain whether Met-1 or Met-49 is the initiator
A:Note: parts of this sequence, including the amino end of the mature protein, were c
R:Itto, N.; Phillips, S.E.V.; Stevens, C.; Ogel, Z.B.; McPherson, M.J.; Keen, J.N.; Y
Nature 350, 87-90, 1991
A:Title: Novel thioether bond revealed by a 1.7 angstrom crystal structure of galacto
A:Reference number: A48244; MUID:91163641; PMID:2002850
A:Contents: annotation
R:Itto, N.; Phillips, S.E.V.; Knowles, P.F.
submitted to the Brookhaven Protein Data Bank, September 1993
A:Reference number: A51740; PDB:1GOF
A:Contents: annotation; X-ray crystallography, 1.7 angstroms, residues 90-728
R:Itto, N.; Phillips, S.E.V.; Knowles, P.F.
submitted to the Brookhaven Protein Data Bank, September 1993
A:Reference number: A51741; PDB:1GOG
A:Contents: annotation; X-ray crystallography, 1.9 angstroms, residues 90-728
R:Itto, N.; Phillips, S.E.V.; Knowles, P.F.
submitted to the Brookhaven Protein Data Bank, September 1993
A:Reference number: A51742; PDB:1GOF
A:Contents: annotation; X-ray crystallography, 2.2 angstroms, residues 90-728
C:Genetics:
A:Gene: gaoc
C:Superfamily: Cladobotryum dendroides galactose oxidase
C:Keywords: disulfide bond; metal binding; oxidoreductase
F:1-64/Domain: signal sequence #status predicted <SIG>
F:55-89/Domain: propeptide #status predicted <PRO>
F:90-728/Product: galactose oxidase #status experimental <MAT>
F:107-116,604-607/Disulfide bonds: #status experimental
F:317-361/Cross-link: cysteinyltyrosine (Cys-Tyr) #status experimental
F:361,584,585,670/Binding site: copper (Tyr, Tyr, His, His) #status experimental
F:419/Binding site: substrate (Arg) #status predicted

Query Match 91.2%; Score 52; DB 1; Length 728;
Best Local Similarity 100.0%; Pred. No. 0.091;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 236 QASSTAPQP 245

RESULT 2

T20823

hypothetical protein R10H10.7 - *Caenorhabditis elegans*C:Species: *Caenorhabditis elegans*

C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jan-2000

C:Accession: T20823; T24156

R:Wild, A. submitted to the EMBL Data Library, April 1996

A:Reference number: Z19329

A:Accession: T20823

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-1416 <W12>

A:Cross-references: EMBL:Z70683; PIDN:CAA94594.1; GSPDB:GN00022; CESP:R10H10.7

R:Barclay, S. submitted to the EMBL Data Library, April 1996

A:Reference number: Z19846

A:Accession: T24156

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-1416 <W12>

A:Cross-references: EMBL:Z70686; PIDN:CAA94614.1; GSPDB:GN00022; CESP:R10H10.7

C:Genetics:

A:Gene: R10H10.7

A:Map position: 4

A:Insertions: 83/3; 254/1; 355/2; 472/3; 548/1; 600/3; 650/1; 733/3; 843/2; 976/3; 1080/3;

C:Superfamily: *Caenorhabditis elegans* hypothetical protein R10H10.7

Query Match

Best Local Similarity 70.2%; Score 40; DB 2; Length 1416;

Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 SSVYAPQPO 11

DB 185 SSVYAPQPO 193

RESULT 3

B75388

NADH dehydrogenase I, C subunit - *Deinococcus radiodurans* (strain R1)C:Species: *Deinococcus radiodurans*

C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 31-Mar-2000

C:Accession: B75388

R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.; S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.

Science 286, 1571-1577, 1999

A:Title: Genome sequence of the radioresistant bacterium *Deinococcus radiodurans* R1.

A:Reference number: A75250; MUID:20036896; PMID:10567266

A:Accession: B75388

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-224 <WMT>

A:Cross-references: GB:AE001994; GB:AE000513; MUID:96459259; PIDN:AAFI1071.1; PID:9645926

C:Genetics:

A:Gene: DR1504

A:Map position: 1

Query Match

Best Local Similarity 68.4%; Score 39; DB 2; Length 224;

Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 SSVYAPQPO 11

DB 79 SSVYAPQPO 87

G75509

hypothetical protein - *Deinococcus radiodurans* (strain R1)C:Species: *Deinococcus radiodurans*

C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 31-Mar-2000

C:Accession: G75509

R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.; S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.

Science 286, 1571-1577, 1999

A:Title: Genome sequence of the radioresistant bacterium *Deinococcus radiodurans* R1.

A:Reference number: A75250; MUID:20036896; PMID:10567266

A:Accession: G75509

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-90 <WMT>

A:Cross-references: GB:AE001910; GB:AE000513; MUID:96458198; PIDN:AAFI0097.1; PID:9645

C:Genetics:

A:Gene: DR0517

A:Map position: 1

Query Match

Best Local Similarity 64.9%; Score 37; DB 2; Length 90;

Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 ASSYAPQPO 11

DB 36 ASSYAPQPO 45

RESULT 5

AH3605

potassium efflux system protein phac [imported] - *Brucella melitensis* (strain 16M)C:Species: *Brucella melitensis*

C>Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 15-Feb-2002

C:Accession: AH3605

R:DeIvecchio, V.G.; Kapral, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanov

Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002

A:Title: The genome sequence of the facultative intracellular pathogen *Brucella melit*

A:Reference number: AD3252; PMID:11756688

A:Accession: AH3605

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-112 <KUR>

A:Cross-references: GB:AE008918; PIDN:AAL54011.1; PID:917984962; GSPDB:GN00191

C:Genetics:

A:Gene: BMEI10769

A:Map position: 11

C:Superfamily: *Pyrococcus abyssi* hypothetical protein PAB1887

Query Match

Best Local Similarity 64.9%; Score 37; DB 2; Length 112;

Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 QASSYAPQPO 11

DB 64 QASSYAPQPO 74

RESULT 6

G87719

protein R119.7 [imported] - *Caenorhabditis elegans*C:Species: *Caenorhabditis elegans*

C>Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 10-May-2001

C:Accession: G87719

R:Anonymous, The C. elegans Sequencing Consortium.

Science 282, 2012-2018, 1998

A:Title: Genome sequence of the nematode *C. elegans*: a platform for investigating bio

A:Reference number: A75000; MUID:99069613; PMID:9851916

A>Note: see websites genome.wustl.edu/gsc/C_elegans/ and www.sanger.ac.uk/Projects/C_

A>Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999;

Mon Jan 13 10:35:14 2003

us-09-554-941-2.fpr

P. J. Ermolaeva, M. J. White, O. J. Saldberg, S. L. Shapiro, L. J. Venter, J. C. Fraser, C. M.
A:Proc. Natl. Acad. Sci. U.S.A. 98, 4186-4191, 2001
A:Reference number: genome sequence of *Yersinia enterocolitica* strain 480/98
A:Accession: AB07372
A:Status: preliminary
A:Residue type: DNA
A:Residues: 1546 <STO>
A:Gene: *Yersinia enterocolitica* strain 480/98
A:Gene: *Yersinia enterocolitica* strain 480/98
A:Superfamily: *Yersinia enterocolitica* strain 480/98

Query Match
Best Local Similarity 63.2% Score 36; DB 1; Length 546;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
OY 1 QASSTAPOP 10
DB 448 QASSTAPOP 457

RESULT 12
TOBPU
C:Species: *Yersinia enterocolitica* strain 480/98
C:Date: 31-Mar-1998
C:Accession: A24746
C:Reference number: genome sequence of *Yersinia enterocolitica* strain 480/98
A:Proc. Natl. Acad. Sci. U.S.A. 95, 5069-5075, 1998
A:Reference number: genome sequence of *Yersinia enterocolitica* strain 480/98
A:Accession: A24746
A:Status: preliminary
A:Residue type: DNA
A:Residues: 1662 <STO>
A:Gene: *Yersinia enterocolitica* strain 480/98
A:Gene: *Yersinia enterocolitica* strain 480/98
A:Superfamily: *Yersinia enterocolitica* strain 480/98

Query Match
Best Local Similarity 63.2% Score 36; DB 1; Length 662;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
OY 2 ASSTAPOP 11
DB 410 ASSTAPOP 419

RESULT 13
hypothetical protein K08E3.1 - *Caenorhabditis elegans*
C:Species: *Caenorhabditis elegans*
C:Date: 15-Oct-1999
C:Accession: T23455
C:Reference number: genome sequence of *Caenorhabditis elegans* strain N2
A:Proc. Natl. Acad. Sci. U.S.A. 96, 4186-4191, 2001
A:Reference number: genome sequence of *Caenorhabditis elegans* strain N2
A:Accession: T23455
A:Status: preliminary
A:Residue type: DNA
A:Residues: 1546 <STO>
A:Gene: *Caenorhabditis elegans* strain N2
A:Gene: *Caenorhabditis elegans* strain N2
A:Superfamily: *Caenorhabditis elegans* strain N2

A:Reference number: 219743
A:Accession: T23455
A:Status: preliminary
A:Residue type: DNA
A:Residues: 1546 <STO>
A:Gene: *Caenorhabditis elegans* strain N2
A:Gene: *Caenorhabditis elegans* strain N2
A:Superfamily: *Caenorhabditis elegans* strain N2

Query Match
Best Local Similarity 63.2% Score 36; DB 2; Length 681;
Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
OY 4 SYTAPOP 10
DB 315 SYTAPOP 321

RESULT 14
Chromatin structural protein homolog Sup5hp - mouse
C:Species: *Mus musculus* (house mouse)
C:Date: 03-Dec-1999
C:Accession: U00004
C:Reference number: genome sequence of *Mus musculus* strain M.129
A:Proc. Natl. Acad. Sci. U.S.A. 96, 4186-4191, 2001
A:Reference number: genome sequence of *Mus musculus* strain M.129
A:Accession: U00004
A:Status: preliminary
A:Residue type: DNA
A:Residues: 1546 <STO>
A:Gene: *Mus musculus* strain M.129
A:Gene: *Mus musculus* strain M.129
A:Superfamily: *Mus musculus* strain M.129

Query Match
Best Local Similarity 63.2% Score 36; DB 2; Length 1082;
Matches 7; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
OY 1 QASSTAPOP 11
DB 883 QASSTAPOP 893

RESULT 15
hypothetical protein XE117 [Imported] - *Xylella fastidiosa* (strain 945C)
C:Species: *Xylella fastidiosa* (strain 945C)
C:Date: 18-Aug-2000
C:Accession: B82722
C:Reference number: genome sequence of *Xylella fastidiosa* strain 945C
A:Proc. Natl. Acad. Sci. U.S.A. 97, 4186-4191, 2000
A:Reference number: genome sequence of *Xylella fastidiosa* strain 945C
A:Accession: B82722
A:Status: preliminary
A:Residue type: DNA
A:Residues: 1546 <STO>
A:Gene: *Xylella fastidiosa* strain 945C
A:Gene: *Xylella fastidiosa* strain 945C
A:Superfamily: *Xylella fastidiosa* strain 945C

A:Reference number: 219743
A:Accession: T23455
A:Status: preliminary
A:Residue type: DNA
A:Residues: 1546 <STO>
A:Gene: *Caenorhabditis elegans* strain N2
A:Gene: *Caenorhabditis elegans* strain N2
A:Superfamily: *Caenorhabditis elegans* strain N2

, F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A.
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawaasak
Aulhairs, da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveira
M.; Tsubako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z
A:Reference number: A59328
A:Contents: annotation
C:Genetics:
A:Gene: XF117

Query Match 61.4%; Score 35; DB 2; Length 483;
Best Local Similarity 63.6%; Pred. NO. 92;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
QY 1 QASSTAPQPQ 11
11: 11111
Db 157 QATGATAPAPQ 167

Search completed: January 13, 2003, 09:57:09
Job time : 13.1538 secs

